CONTEMPORARY SCIENCE EDUCATION RESEARCH: LEARNING AND ASSESSMENT

A collection of papers presented at ESERA 2009 Conference

EDITORS

G. ÇAKMAKCI & M. F. TAŞAR

CONTEMPORARY SCIENCE EDUCATION RESEARCH:

LEARNING AND ASSESSMENT

CONTEMPORARY
SCIENCE EDUCATION RESEARCH:
LEARNING AND ASSESSMENT

Edited by

GÜLTEKİN ÇAKMAKCI
Hacettepe Üniversitesi, Ankara, TURKEY

MEHMET FATİH TAŞAR
Gazi Üniversitesi, Ankara, TURKEY

PEGEM AKADEMİ

Istanbul ESERA 2009
European Science Education Research Association

ISBN - 978-605-364-033-2

Referencing articles in this book

The appropriate APA style referencing of articles in this book is as follows:

Chin, C., & Osborne, J. (2010). Students' questions and discursive interaction: how they impact argumentation during collaborative group discussions in science. In G. Cakmakci & M.F. Taşar (Eds.), *Contemporary science education research: learning and assessment* (pp. 3–12). Ankara, Turkey: Pegem Akademi.

TABLE OF CONTENTS

PART 2: Assessment of Student Learning and Development

Preface

At the ESERA 2010 Conference, over 1000 researchers met in Istanbul to exchange experiences and discuss contemporary issues in science education. This book is a part of series of the proceedings of the ESERA 2010 Conference. The book consists of two parts. The first part includes 31 papers on theories, strategies, and models of learning and the second part includes 29 papers on assessment of student learning and development.

We wish to thank all of the contributors in this book for their hard work

G. Çakmakcı
Hacettepe Üniversitesi, Ankara

M. Fatih Taşar
Gazi Üniversitesi, Ankara

PART 1
LEARNING SCIENCE

STUDENTS' QUESTIONS AND DISCURSIVE INTERACTION: HOW THEY IMPACT ARGUMENTATION DURING COLLABORATIVE GROUP DISCUSSIONS IN SCIENCE

Christine Chin
Nanyang Technological University, Singapore

Jonathan Osborne
Stanford University, USA

Abstract

This study investigated the potential of students' questions both as an epistemic probe and heuristic for initiating collaborative argumentation in science. Four classes of students, aged 12-14 years from two countries, were asked to discuss which of two graphs best represented the change in temperature as ice was heated to steam. The discussion was initiated by asking questions about the phenomenon. Working in groups and guided by a set of question prompts, an argument sheet, and an argument diagram (with members who had different viewpoints), different arguments were discussed and one group of students from each class audiotaped. The number of questions written, the concepts addressed, and the quality of written arguments were then scored. A positive correlation between these factors was found. Discourse analysis showed that the questions prompted students to articulate their puzzlement; make explicit their claims and (mis)conceptions for peers to respond to; identify and relate relevant key concepts; construct explanations; and consider alternative propositions when their ideas were challenged. Productive argumentation was characterized by questions which focused on key ideas of inquiry and a variety of scientific concepts, and explicit reference to the structural components of an argument. Implications for instruction are discussed.

Introduction

Argumentation is believed to play an important role in students' learning of science. First, argumentation is central to the process of thinking and scientific reasoning, and plays an important role in developing conceptual understanding. Second, discursive activities that involve students in argumentation present a more authentic image of the nature and practice of science (Duschl & Osborne, 2000; Newton, Driver, & Osborne, 1999). To engage students in the reasoning and discursive practices of scientists, a key focus of research in this area has been the design of learning environments that attempt to incorporate and promote argumentation in science (e.g. Osborne, Erduran, & Simon, 2004a; Zohar & Nemet, 2002). Since argumentation is a social activity where ideas are explored through dialogue, invariably these designs have involved collaborative group discussions that engage students in epistemic practices such as proposing, justifying, evaluating and criticizing knowledge claims (Jiménez-Aleixandre, 2008).

Other studies have evaluated the quality of students' arguments by examining the form and content of the arguments (e.g., Clark & Sampson, 2008; Sandoval & Millwood, 2005). The main findings are that students often have difficulty constructing arguments on their own although their argumentation skills can be enhanced by various forms of intervention, such as explicit teaching or the provision of conceptual scaffolds. These conceptual scaffolds may include the use of diagramming (e.g., Schwarz, Neuman, Gil, and Ilya, 2003; van Amelsvoort, Andriessen, and Kanselaar, 2007) that visually portray the components of an argument being discussed. Yet another way of supporting argumentation is to develop instructional scaffolds as part of the participant structure and social norms of collaborative debate. For example, Herrenkohl and Guerra (1998) and Herrenkohl, Palincsar, deWater, and Kawasaki (1999) designed a questions chart to help students ask questions about predictions, theories, and results.

Previous work on collaborative discourse and argumentation has focused mainly on the structure of the overall activity such as the group organization and ground rules for procedures. Little is known, however, about the specific discourse moves involved in the constructive processing of ideas although some studies (e.g. Kim and Song, 2006) have suggested that questioning plays a pivotal role in having the potential to be used both as an epistemic probe *and* heuristic tool for initiating argumentation in inquiry-based science. Indeed, Hogan, Nastasi, & Pressley (1999) found that effective groups (in comparison to the ineffective ones) often questioned each other and the questions initiated sustained episodes of knowledge construction. Questions can coordinate group interactions by keeping the members focused on the task at hand, exposing students' thinking and making it available for discussion, activating prior knowledge, clarifying doubts, eliciting explanations, justifying reasoning, stimulating different kinds of thinking, guiding cognitive processing, and challenging diverging viewpoints – all of which contribute to the subsequent negotiation of understanding (Chin & Osborne, 2008). Students could thus be encouraged to ask a variety of "critical" or other questions in an argumentative task which allow the interlocutor or respondent to probe the various aspects of an argument or to challenge and object to an argument.

It is suggested that supporting students in question-asking could stimulate them to engage in meaningful argumentation. This study, therefore, created the conditions for collaborative discourse that focuses on questioning *and* argumentation, by having students pose questions and respond to each other when they were put together in small groups to discuss opposing viewpoints. The study aimed to explore (a) the relationship between students' questions (in terms of the number of questions and relevant concepts addressed by the questions) and the quality of the arguments generated when students worked in small groups to resolve conflicting ideas about a scientific phenomenon; and (b) how students' questions contributed to the argumentative process during collaborative group discussions.

Methods

Four classes of students (n=129), aged 12-14 years, participated in this study. Two classes (years 8 and 9) were from two schools (E1 and E2) in London, England, while the other two classes (years 7 and 8) came from two other schools (S1 and S2) in Singapore. A criterion for selecting the classes was that the students needed to have some prior knowledge of the particle model of matter in the solid, liquid and gaseous states for water. Yet, they should not have undertaken any activity to measure the variation in temperature as water changes state from ice to water vapour, or had any explanations of the phenomenon provided. Otherwise, there would be little for them to discuss or argue about. Data-collection for each class was conducted over one session. The students were briefed initially by their teachers on the structure of an argument and then on the importance of asking questions to ascertain the information required for making claims, evaluating arguments, and making decisions.

The students then worked in groups (with members having different viewpoints) on a given task where they had to decide which of two given graphs (A or B) best represented how the temperature would change as they heated ice to steam. This task was adapted from Osborne, Erduran, and Simon (2004b) and a list of evidence statements, such as "Ice will melt when it is heated and turns into water" and "Ice melts at 0^{o}C and boils at 100 ^{o}C", was provided. Students were encouraged to ask questions based on the given task and were given a list of question prompts such as "What pattern or trend do I see here?", "What is my explanation for how this happens?", and "What is the evidence to support my view?" to scaffold the question-asking process. The students first individually wrote questions about the graphs, then took turns to pose their questions and recorded their pooled questions onto a Question Web. The latter was a graphic organizer resembling a spidergram where students listed their questions.

Using the Question Web and question prompts provided, students asked questions of each other and used the given evidence statements, as well as their other ideas, to justify why they believed their choice of graph was correct. To help them structure their arguments, students were given an "Our argument" sheet containing writing stems which required them to state their claim, data or evidence, reason(s) for their answer, counter-argument and rebuttal. To help students visualize their ideas diagrammatically, each group was also required to represent their argument in graphic form using an Argument diagram, which was based on Toulmin's (1958) structure of an argument. One group of students from each class was selected for audio-taping and the taped group discourse transcribed verbatim. The criteria for selection were that the chosen students should be (a) verbally expressive so that the discourse could be captured on tape; (b) likely to stay on-task; (c) able to work with each other cordially; and (d) hold opposing viewpoints on the answer in the given task.

The *written* questions from all the Question webs of the 29 groups were classified by question type and counted. The questions asked by each group were then scored for the number of distinct scientific concepts (e.g. boiling point of water, constant temperature at 0°C) associated with the given problem. To determine the quality of the written argument generated, each group's responses on both the Argument sheet and Argument diagram were analyzed. One point was awarded for each well-reasoned and explicitly stated component of the argument that pertained to the claim, a piece of data or evidence, a warrant, elaboration of background ideas (backing), a qualifier, a counter-argument, or a rebuttal. In addition, one extra point was given for the correct answer (i.e., graph B). No points were given for questions without the relevant content. The sum of the scores across the various components of an argument provided an "argument score" for the group, which represented the quality of the argument generated. The maximum number of points possible for an exemplary "model" argument was 11, which included the correct answer for the claim. The inter-rater reliability in scoring between the first author and a second scorer (a science educator who did not have a stake in this study and who was blind to the countries or schools from which the students came) was 97% after resolving differences through discussion. The number of questions written, as well as the number of relevant concepts addressed by these questions, was then correlated with the argument scores, using SPSS.

To determine how students' *oral* questions contributed to the dialogic argumentation within a group, the verbal discourse from each of the four target groups was examined to explore how the questions influenced the evolution of students' ideas and reasoning about the concepts relevant to the graphs. After segmenting the transcript of each group discussion into argumentative episodes, the role of the questions posed and the thinking underlying each utterance was then inferred. Discourse analysis focused on both the function of the questions and the content of the talk that followed.

Results

Relationship between the Questions and Arguments Written by the Students

The emergent categories of questions posed by the students on their Question webs are referred to as: (a) key inquiry; (b) basic information; (c) unknown or missing information; (d) conditions under which the heating was carried out; and (e) others. "Key inquiry" questions were mainly "Why?" questions that sought *explanations* for the shape of the graph, the flat sections corresponding to 0°C and 100°C, the differences in the rate of temperature change, why the temperature decreased below 0°C for ice, and why the temperature rose above 100°C for steam. They addressed the big but fundamental ideas related to the phenomenon of heating ice to steam, and had the potential to address the salient concepts necessary to construct a coherent argument.

"Basic information" questions addressed the most basic, *factual information* that students needed to know before they could answer the key inquiry questions. They pertained to the melting point of ice, the boiling point of water, the properties of a graph, and the particle theory of matter. Questions that identified any *information not given* in the task sheet and which students felt were necessary for them to arrive at an informed answer, constituted a category named "unknown or missing information". Such questions pertained to the nature of the heat supply, whether impurities were present, the surroundings in which the heating was carried out, the room temperature, the volume of ice used, and the surface area of the container. Questions in the "conditions" category showed *predictive thinking* in that students asked what would happen if the conditions under which heating occurred were altered. They included speculations about changes in the intensity of the flame, ambient temperature, amount of ice used, and time of heating. Finally, miscellaneous questions were placed in a category labelled "others". They included questions dealing with the validity and relevance of the given data.

Table 1 shows the distribution by number and type of questions for all the student groups. The groups asked an average of 6 questions each. Most were key inquiry (42.9%) and basic information (22.0%) questions. A mean of 5 different concepts were addressed by the questions from each group. The mean argument score was 7 (s.d.=2.53) out of a maximum possible of 11. Figure 1 shows the argument written by Group 3 on the Argument sheet. It is given as an example of how the students' written arguments were evaluated. This argument scored 10 points.

Table 1. Distribution of written questions by number and type for all the student groups

Type of question	No. of questions	Percentage of total
Key inquiry	72	42.9
Basic information	37	22.0
Missing or unknown information	24	14.3
Conditions	10	5.9
Others	25	15.9
Total	168	100.0

There was a moderate, positive correlation between the total number of questions asked and the argument score (r_1 = 0.38, p=.045). However, the correlation between the *number of concepts addressed by the questions* and the argument score (r_2 = 0.51, p=.006) was higher. Since most (42.9%) of the questions posed were "key inquiry" ones, this shows that when more questions focusing on the key scientific concepts were asked, the arguments that were generated were of a better quality. More significant, perhaps, was the finding that while the total number of questions asked played a role in determining the quality of the argument generated, it was the *type* of questions posed that mattered more. Those questions that targeted the key inquiry concepts that were fundamental to scientific understanding had the greatest potential to contribute to elaborate and higher-quality arguments.

Contribution of Students' Questions to Dialogic Argumentation During Group Discussion

The questions generated by the students played an important role in the argumentative process. In particular, the questions:

a) foregrounded any incomprehension or points of disagreement students had; and pushed them to articulate these either covertly to themselves or overtly to their peers.
b) prompted students to make explicit their beliefs and claims, thus displaying their (mis)conceptions for peers to examine.
c) spurred students to identify the key concepts relevant to the scientific phenomenon at hand.
d) impelled students to make connections between their various ideas, and elicited self- or peer-explanations of the scientific phenomenon.
e) challenged opposing viewpoints, stimulating peers to critically evaluate their initial ideas and to consider alternative propositions.

Examples of how students' questions contributed to each of the five above-mentioned areas are given below. Excerpts of discourse are taken from the group in school S2, which comprised Amy (A), Devi (D), and Val (V). Prior to the group discussion, both Amy and Val chose graph B but Devi believed that neither graph was correct.

a) Pushing students to be aware of and to articulate their puzzlement.

The following excerpt is taken from the beginning of group GS2's discussion when the students were brainstorming questions about the given graphs. Amy was not sure whether the temperature of steam could rise above 100°C. Val then responded that 100°C was the boiling point and that it was possible for the temperature of steam to go beyond that temperature.

A: *When steam, when the water becomes steam, right, is the temperature more than 100°C? Will the temperature go more than 100°C?*
V: Yes. I think it can.
A: *It boils at 100°C, right? Does steam... go above 100°C?* No. It's not water.
V: It's water. Water at boiling point is 100°C.
A: Yeah, but it might not go above 100°C.
V: It can […] because it turned into steam already. You know, it changed. 100°C is only the boiling point.
A: *This temperature is for the steam or for the water?* For the water, right?.... *Can the temperature of steam go higher than 100°C?*
V: Yes, it can, because 100°C is just the boiling point of water. But steam can be [at] a higher temperature than 100°C.

In thinking about the questions that she had, Amy became aware of her uncertainty about the temperature of steam. Verbalizing her questions aloud led Val to respond to and clarify her doubts. Val's initial assertion in the second turn ("Yes, I think it can"), that the temperature of steam could exceed 100°C comprised a simple claim. However, her subsequent response to Amy in the sixth turn ("It can because it turned into steam already…. 100°C is only the boiling point") contained a simple justification with data and warrant.

b) Prompting students to make explicit their beliefs, claims, and (mis)conceptions.

This next excerpt shows the students continuing to discuss their ideas about the boiling of water to steam, as well as explicitly stating their choice of graph. Devi's misconception about the shape of the graph at the melting point of ice was made visible in the fourth turn below when she said that "the temperature should constantly rise and not stop at 0°C even though it's the melting point".

A: *What is the temperature of steam? How high can it go beyond 100°C? And does water continue to boil after it reaches 100°C?* We are not very sure of the answer but we feel that water does continue to boil as it reaches 100°C. ….
D: We all feel that graph A is wrong....
V: And that graph B is correct!
D: I, however, feel that neither are [is] right because I believe that before 0°C, energy is already collected and the temperature should constantly rise and not stop at 0°C even though it's the melting point. I feel that the temperature should continue rising because it's still melting. And as it rises, it stops at 100°C because... the boiling temperature is still 100°C where the water will evaporate... Then the steam is above 100°C because steam does not have a fixed temperature.
V: I disagree.
A: Both Val and I… feel that graph B is the correct graph because the temperature does rise but when the temperature reaches 0°C, it is constant for a while. It is constant while the ice changes from a solid state to a liquid state. And [this] is when the energy is used to break the bonds. And thus, there will be no temperature change. That's why the temperature is constant at 0°C. However, it continues to rise and when it reaches 100°C, it is constant again for another period of time. And that is when the water changes from the liquid state to the gaseous state. And energy is also being used then to break the bonds between the particles and thus, the temperature remains constant. So what we're saying is […] The ice takes time when it reaches 0°C to melt. So the temperature remains constant because the energy has to be collected at that particular temperature.
D: Yes. Okay, I admit. Alright. I feel that maybe your answer might be correct.

c) Spurring students to identify key concepts related to the scientific phenomenon.

In this example, the key concept being identified was the conditions necessary for graph B to be valid. This addressed the qualifier in the argument. While working on the Argument diagram which required them to think about the qualifier, the students considered the various possible conditions under which graph B would be tenable.

V: Qualifier. *What are the conditions?....*
A: Heat source. State of ice.
D: Temperature of ice.
V: Uhm, intensity of the flame. *What else?*
A: That no salt is added to it….
V: Uh, amount of ice. Surface area of the beaker. Surface area.
A: How the ice is heated. One might be heated in a long, narrow beaker, whereas the other could be heated in a shallow dish. And this would affect the reading.
V: Yeah, because the surface area can affect the rate of evaporation....
A: .... The temperature of the ice, the intensity of the flame for heating, and the surface area of the ice exposed to the heat.... So in our conditions, we also have to make sure that the heat is […] the supply is constant and evenly distributed.... So you can't argue that the heat supply is […] you know, unevenly distributed.

In having to answer the question pertaining to the conditions under which graph B would be supported, the students were prompted to think about the necessity of having pure water where there would be no impurity such as salt, the intensity of the heat source, and the way the ice was heated.

d) Impelling students to make connections between ideas and eliciting explanations.

In the excerpt below, the students were grappling with their puzzlement about why there were horizontal plateaus on graph B at 0^oC and 100^oC. Val, who believed in graph B, first posed the question of why the temperature remained constant at those points. Then she generated a self-explanation where she answered her own question. In trying to make sense of the disparate bits of information presented in the given graphs and evidence statements, she made links between them and consolidated them into her argument. She synthesized the various pieces of information related to the changes of state of water, which also addressed the different components of an argument. These concepts addressed the data (viz., "ice melts at 0^oC and boils at 100^oC") and backing (viz., "while energy is being used to break bonds between particles, there will be no temperature change"). Subsequently, when Devi (who initially thought that neither graph A nor B was appropriate and drew her own graph C which did not show a flat portion at 0^oC) asked why the temperature could not just rise continuously without a pause, Val explained that the heat energy would be converted to kinetic energy during the change of state.

V: My question for graph B is '*Why is the temperature constant at 0^oC and at 100^oC only, and not at other temperatures?*'. Because I think… energy is being used to break bonds between particles, then there will be no temperature change. And also ice melts at 0^oC and boils at 100^oC. So during that time when ice melts and changes to water, the bonds are broken and that is why there is no temperature change at that time because energy is being used. And also when it changes to steam at 100^oC....
D: I came up with graph C and it does not pause at 0^oC.... *Isn't the heat energy enough […] collected enough before it reaches 0^oC? So can't it [temperature] just go all the way up constantly to 100^oC?*
V: …While energy is being used to break bonds between two particles, there will be no temperature change. So at 0^oC, ice is melting so that it turns from the state of solid to the state of liquid. So heat energy is being used to be converted to kinetic energy where the molten molecules are vibrating to move… to overcome their forces of attraction and to become liquid molecules.
D: *But how come when it reaches 0^oC, it takes quite a long time for it to start changing its temperature? Couldn't it start melting straight away*?
A: That also depends on the division of the time on the graph…. Because in both graphs A and B…. if the timings are different, then that would greatly affect our argument….

When Devi queried why the temperature had to stay "quite a long time" at 0^oC before it increased again, Amy pointed out that the scale for the time on the horizontal axes of the graphs was not given. After a few rounds of negotiating their ideas, Devi finally changed her mind to accept that graph B would be more appropriate.

e) Challenging opposing viewpoints, stimulating critical evaluation of initial ideas, and prompting consideration of alternative propositions.

The following excerpt shows the conversation when the students were responding to the section on their Argument sheet which asked for a possible counter-argument and rebuttal.

V: Someone might argue against our idea by saying that energy does not require so much time to break the bonds, especially since the time is not stated on the graph…. Like for example, if someone says that ice can straightaway melt and turn into the liquid state, we would say that energy is being used to break the bonds between the particles. So time is needed to change its state.
D: The temperature stays constant at 0^oC for a while.
A: Someone can also state that the heat source is.... very, very strong, so causing it to melt instantly and its temperature to constantly rise up instead of stopping and pausing for a while.
V: *So how are you going to rebut?*
A: Uhm […] To rebut that, we'd probably have to carry out an experiment where all the factors are constant and we can try to use a constant heat source in one experiment and another not so constant….
V: We can use different kinds of flames [of] different intensity to observe. And see what happens.
D: So we'll show her an experiment.

The students considered the situation when someone might propose the counter-argument that if the heat source were "very, very strong", the ice would then "melt instantly". Thus, the time taken for melting would be so short as to be negligible. In such a case, the flat portion on the graph would then be so insignificant as not to be registered at all.

We see that students' questions can function in five different ways (listed above from a to e) in contributing to the process of argumentation. The questions that students posed to each other elicited ideas that nudged them to address the different components of an argument. For example, in verbalizing their questions aloud, the students became aware of what they did not understand or were unsure of. In responding to questions that required them to make explicit their beliefs, the students stated their *claims*. Questions that interrogated fundamental concepts, unknown information, assumptions, and conditions led students to think about the *data, evidence, backing*, and *qualifier* involved in an argument. Having to respond to questions that required them to make connections and generate explanations or reasons, pushed students to address the *warrant* in the argument that they were constructing. And finally, in challenging viewpoints that were different from their own, the students put forward their *counter-arguments* and *rebuttals*.

Discussion and Conclusions

This study has shown that it is possible to engage students in a discourse of productive argumentation mediated through the use of their own questions, and the preceding extracts give an idea of the range of roles that these questions play. The Argument sheet and argument diagram, together with the list of question prompts and evidence statements, encouraged students to talk about their ideas related to the components of an argument. There was a moderate, positive relationship between the questions that students asked and the quality of their argument. The finding that questions which addressed the substantive scientific concepts contributed to more elaborate and higher-quality arguments is reminiscent of that by Harper, Etkina & Lin (2003) who found that it was the *type* of questions asked, rather than the number, that was associated with better conceptual achievement in physics.

Why do questions help students to engage in dialogic argumentation? Students' questions served as a thinking device to foster critical dialogue, both when the students were working towards consensus of ideas and when resolving opposing views. The act of individually brainstorming and then verbalizing the questions aloud made the students aware not only of what they understood but also what they did not understand about the given graphs (excerpt a). Publicly articulating their puzzlement to each other prompted them to make explicit their beliefs, claims, and (mis)conceptions which their peers could respond to (excerpt b) and enabled them to formulate objections or counter-arguments by the meta-linguistic moves of supporting the other side of the question, questioning the truth of a claim, or criticizing the reasoning supplied (excerpt e) (Leitão, 2000). The questions helped to heighten students' awareness of the salient concepts related to the changes of state of water, impelled them to make connections between their ideas, and elicited both self- and peer- explanations. When students had divergent viewpoints, they also used questions to challenge their peer's thinking and this prompted the critical evaluation of initial ideas as well as the consideration of alternative propositions. Facilitating such dialogic interaction is important as the interplay between construction and critique is the hallmark of scientific thinking by which students:

> come to *know that* scientific knowledge is held accountable by explicit connections to nature, to *know how* to play the roles of constructor and critiquer appropriately, and to *know that* the interaction of these roles in practice yields reliable knowledge. Ford, 2008 (p405, author's emphasis)

One of the most significant contributions of the questions was its potential in scaffolding students' argument construction by eliciting the epistemic features of explanations. Students' questions stimulated the co-elaboration and justification of various positions as several points of view were examined and modified to produce a new idea that took into account the different standpoints. They served as triggers to enable argumentative and epistemic moves, such as concessions, challenges and counter-challenges, which subsequently led to the construction of more elaborate explanations and justifications, as well as to changes in the standpoints of members who modified their initial conceptions. To facilitate such dialogic talk, however, students need to be supported by a variety of scaffolding structures. Specifically:

- Scaffolds need to be provided for students to ask questions.
- Students need to be taught the structural components of an argument (including the appropriate vocabulary), criteria for a good argument, and the language of argumentation.
- Argument diagrams provide a structure to help students focus, organize, and verbalize their arguments both orally and visually.
- A participant structure that encourages students to pose questions, make justified claims, construct explanations, and challenge opposing viewpoints is required.
- Students need to work towards a consolidated and consensual written product that is a solution to the scientific problem that they are working on.

Helping students learn to ask the right kinds of questions and build onto each other's thinking may be a key to orchestrating argumentation. Future research could examine whether an extended programme that involves the explicit teaching of questioning and argumentation skills, together with multiple opportunities for practice in different contexts, would lead to enhanced abilities in student argumentation.

References

Chin, C. & Osborne, J. (2008). Students' questions: a potential resource for teaching and learning science. *Studies in Science Education*, 44(1), 1-39.

Clark, D. B. & Sampson, V. (2008). Assessing dialogic argumentation in online environments to relate structure, grounds, and conceptual quality. *Journal of Research in Science Teaching*, 45(3), 293-321.

Duschl, R.A. & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education*, 38, 39-72.

Ford, M. (2008). Disciplinary authority and accountability in scientific practice and learning. *Science Education, 92*(3), 404-423.

Harper, K.A., Etkina, E., & Lin, Y. (2003). Encouraging and analyzing student questions in a large physics course: Meaningful patterns for instructors. *Journal of Research in Science Teaching, 40*(8), 776–791.

Herrenkohl, L. R., & Guerra, M. R. (1998). Participant structures, scientific discourse, and student engagement in fourth grade. *Cognition and Instruction, 16*, 431–473.

Herrenkohl, L.R., Palincsar, A.S., deWater,L.S. & Kawasaki, K. (1999). Developing scientific communities in classrooms: A sociocognitive approach. *Journal of the Learning Sciences*, 8(3&4), 451-493.

Hogan, K., Nastasi, B. K., & Pressley, M. (1999). Discourse patterns and collaborative scientific reasoning in peer and teacher-guided discussions. *Cognition and Instruction, 17*, 379–432.

Jiménez-Aleixandre, M.P. (2008). Designing argumentation learning environments. In S. Erduran & M.P. Jiménez-Aleixandre (Eds.). *Argumentation in science education: Perspectives from classroom-based research* (pp.91-115). Dordrecht, The Netherlands: Springer.

Kim, H. & Song, J. (2006). The features of peer argumentation in middle school students' scientific inquiry. *Research in Science Education*, 36(3), 211-233.

Leitão, S. (2000). The potential of argument in knowledge building. *Human Development, 43*, 332-360.

Newton, P., Driver, R., & Osborne, J. (1999). The place of argument in the pedagogy of school science. *International Journal of Science Education*, 21, 553-576.

Osborne, J., Erduran, S., & Simon, S. (2004a). Enhancing the quality of argumentation in science classrooms. *Journal of Research in Science Teaching*, 41(10), 994 – 1020.

Osborne, J., Erduran, S., & Simon, S. (2004b). *Ideas, evidence and argument in science.* In-service training pack, resource pack and video. London: Nuffield Foundation.

Sandoval,W. A.,& Millwood, K. (2005). The quality of students' use of evidence in written scientific explanations. *Cognition and Instruction*, 23(1), 23 – 55.

Schwarz, B.B., Neuman, Y., Gil, J. & Ilya, M. (2003). Construction of collective and individual knowledge in argumentative activity. *The Journal of the Learning Sciences*, 12(2), 219-256.

Toulmin, S. (1958). *The uses of argument.* Cambridge, UK: Cambridge University Press.

van Amelsvoort, M., Andriessen, J., & Kanselaar, G. (2007). Representational tools in computer-supported collaborative argumentation-based learning: How dyads work with constructed and inspected argumentative diagrams. *Journal of the Learning Sciences, 16*(4), 485 - 521.

Zohar, A., & Nemet, F. (2002). Fostering students' knowledge and argumentation skills through dilemmas in human genetics. *Journal of Research in Science Teaching*, 39(1), 35– 62.

Figure 1. Argument written by Group 3 on their Argument sheet.

Problem:

Which graph is most likely to show how the temperature of water changes as it heats up?

	Group's written response	Score
Our claim	We think that the graph most likely to show how the temperature of water changes as it heats up is ….. Graph B.	1
Data / Evidence	Our evidence for this is ….. Graph B stays at the same place at 0°C and 100°C.	1+1
Reason (Backing) (Warrant)	This evidence supports our idea because ….. In solids, there are bonds between the particles that hold them together in a fixed shape. Ice melts at 0°C and [water] boils at 100°C. Energy is needed to break bonds between particles. While energy is being used for this, there is no temperature change. We do not think that Graph A is correct because ….. It [the line graph] moves at a steady speed and does not stop to break the bonds between the particles like in Graph B.	1 1 1 1
Counter-argument	Someone might argue against our idea by saying that ….. When you heat a substance, the supply of heat is usually constant.	1
Rebuttal	If someone does not agree with us, we would convince him / her by ….. Showing an ice cube melt and see whether it [the line graph] moves at a steady speed or not.	1
	Argument score Total argument score (+ 1 for correct answer)	9 10

Note:

- Points were awarded for each component of the argument as shown above.
- The maximum possible total score was 10 + 1 (for the correct answer, Graph B) = 11.
- This group failed to score 1 point for an appropriate qualifier (e.g. that the water was pure and did not contain any impurities, which would then lower the melting point below 0°C and raise the boiling point above 100°C).

SOURCES OF OPTICS MISCONCEPTIONS

Derya Kaltakci
Kocaeli University

Ali Eryilmaz
Middle East Technical University

Abstract

Starting from early 1970s students' conceptual understanding on many physics concepts have been studied by a variety of researchers in physics education research. A practical goal of this research is to identify appropriate methodologies to eliminate these unscientific conceptions, as well as to understand the nature of acquisition of scientific conceptions fully. However, the sources of the unscientific conceptions and the elimination of these sources have been given little importance by researchers. Students' experiences, textbooks, language used, teachers can be listed among several possible sources of students' unscientific conceptions in physics. Optics in physics is one of the topics in which students have many unscientific conceptions. In the present study examples of these possible sources will be provided with related optics misconceptions.

Introduction

In physics education research, studies on students' difficulties in conceptualization of scientific concepts and their misconceptions have gained impetus and constituted the topic of many studies in this field since 1970s. The pioneering attempts of Professor Lillian McDermott and Physics Education Group of University of Washington have provided invaluable insights into student difficulties and illustrate how research can guide the design of research based instructional materials for addressing student difficulties. Nowadays, a common application of physics education research is to use appropriate methodologies to address these unscientific conceptions. However, the sources of the unscientific conceptions and the elimination of these sources are given little importance by researchers.

Al-Rubayea (1996) categorized sources of students' answers into three groups in a misconception test as: the sources of all answers, the sources of correct answers, and the sources of incorrect answers. The most frequent sources for incorrect ideas of Saudi Arabia students were listed as guessing, physics books, general information, teacher, experience, TV or radio, newspaper or magazine, the phrasing of the question itself, and parents. Ivowi (1984) stated that teachers and textbooks were the primary sources of Nigerian students' misconceptions. Heller and Finely (1992) identified teacher misconceptions as a probable source of students' misconceptions. Beaty (1987) claimed that students learn misconceptions from physics textbooks since misconceptions are presented as facts in students' textbooks. Iona (1987) stated that if the textbooks had fewer errors, some misconceptions might not be widely distributed or gain acceptance.

Optics in physics is one of the topics in which students have a number of unscientific conceptions and should be studied in detail to determine the sources of these conceptions. Students' experiences, textbooks, language used, teachers can be listed among several possible sources of students' unscientific conceptions in optics. In the present study, examples of these possible sources will be provided together with related optics misconceptions.

Rationale

In educational research studies, not only the identification of misconceptions in optics is crucial, but also the identification of possible sources of these misconceptions for elimination or refinement of these possible sources is important. For this reason there is a necessity of careful investigation of possible sources of student misconceptions in optics. With this aim, the major purpose of the present study is to provide examples of possible sources of optics related misconceptions, to take the attention of authorities on these possible sources.

Methods

In this study document analysis method was used in order to analyze the optics related articles and sections of physics textbooks. The use of documents often entails a specialized analytic approach called content analysis (Marshall & Rossman, 1999). Content analysis is a technique which enables researchers to study human behavior in an indirect way through their communication (Fraenkel & Wallen, 2000). In order to determine possible sources of optics related misconceptions, a large number of articles and textbooks are analyzed carefully by content analysis technique and the most striking examples are provided in this study.

Results

With careful study of articles in literature and the textbooks with content analysis, some possible sources of optics misconceptions are discussed and exemplified in this section.

Sources of Misconceptions

Personal Experience

The close relation of optics topics to the daily experiences of students' may be a source of misconceptions in this topic. For example; many students have the belief that moving back from a vertically held plane mirror will allow them to see more of the image of their body. In reality, however, the student sees the same amount of his body regardless of distance from the mirror. The size of the mirror relative to the observer, not the distance from the mirror is critical in how much of the image of your body can be seen in plane mirrors. Goldberg and McDermott (1986) postulated that the use of mirrors in daily life may responsible for this misconception. In daily life, we use small mirrors to see our face, and big mirrors to see whole body. By standing farther back from the mirror we can see more of our body with a minimum eye movement. This experience with plane mirrors may contribute to this misconception.

Textbooks

Textbooks have an important role in students' construction of scientific understanding. However, sometimes textbooks may responsible for student misconceptions, since they may present incomplete arguments or misleading statements. Some may also contain incorrect statements. As an example, many students have the misconception that only the lens of the eye is responsible for the refraction of light. In many physics and science textbooks, human eye is discussed as an optical device; however, in representations including the light rays entering the eye, light rays are illustrated to bend at the lens of the eye only as illustrated in Figure 1. However, in reality almost 70 per cent of the refraction occurs at the cornea of the eye where the refractive index changes the most as in Figure 2. These representations in textbooks may be considered as a possible source of this misconception.

Language

Language used by individuals may sometimes responsible for student misconceptions. For instance, many students have the misconception that color is a property of objects rather than light. In daily language saying "the table is red" instead of "the table is reflecting red light" may be considered as the source of this misconception. In the second sentence we emphasize the red color to be the property of light, whereas in the first sentence as the property of table.

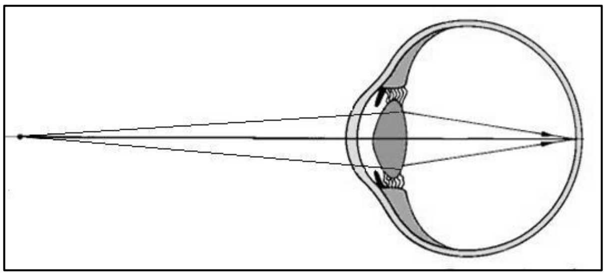


Figure 1. Example of Incomplete Representation of Refraction of Light Rays on the Eye.

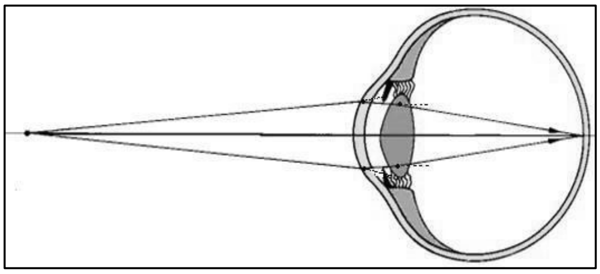


Figure 2. Example of Complete Representation of Refraction of Light Rays on the Eye.

Teachers

Teachers have an important role in students' conceptions of scientific concepts. In some cases teachers may be unaware of student difficulties and thus fail to take appropriate care in presenting certain ideas to students. Moreover, if teachers have misconceptions themselves, they may transfer them to the students. Studies conducted on teachers (Bendall, Goldberg and Galili, 1993; Feher and Rice, 1987; Galili, Bendall and Goldberg, 1993; Heywood, 2005; McDermott, Heron, Shaffer and Stetzer, 2006) have revealed that teachers have misconceptions on several concepts in optics. For example, prospective teachers in the studies of Feher and Rice (1987) and Bendall, Goldberg and Galili (1993) had the conception about shadow as the presence of something, rather than as the absence of light. With teachers having such a misconception, it may not be surprising for their students to have such a misconception.

Conclusions and Suggestions

In conclusion, there are many possible sources of student misconceptions. In many cases there is evidence that these ideas are deeply held and resist to instruction. None-the-less it would make sense to minimize the contributing factors that might help promote the development of these incorrect ideas.

Authors of the textbooks could take care to take into account known student problems in writing science textbooks. Teachers as the users of these textbooks should be on the alert about the information presented with the textbooks since they do not always contain the scientifically accepted knowledge. Also teachers should be watchful in using language correctly in and out of class. If there is a mismatch between science conceptions and everyday experiences, then students should be confronted with them and these misconceptions should be refined with their scientific ones.

The teachers have a great role in this process in identification and elimination of optics related misconceptions, for this reason, in teacher education process there should be given importance to the discussions about optics related misconceptions and the possible sources of these misconceptions. Also there should be given great care to ensure that teachers have a deep subject-matter background in the topics they are to teach.

Acknowledgements

We thank Peter S. Shaffer for providing substantial feedback on this paper. We also appreciate the ongoing financial support of the Scientific and Technical Research Council of Turkey and Faculty Development Programme at Middle East Technical University.

References

Al-Rubayea, A. M. (1996). *An Analysis of Saudi Arabian High School Students' Misconceptions About Physics Concepts.* Unpublished doctoral dissertation, Kansas State University, Manhattan, Kansas.

Beaty, W. J. (1987). The origin of misconceptions in optics. *American Journal of Physics,* **55**, 872-873.

Bendall, S., Goldberg, F. & Galili, I. (1993). Prospective teachers' prior knowledge about light. *Journal of Research in Science Teaching,* **30**(9), 1169-1187.

Feher, E. & Rice, K. (1987). A comparison of teacher-student conceptions in optics. In J. D. Novak (Ed.), Proceedings of the Second International Seminar: Misconceptions and Educational Strategies in Science and Mathematics, (pp. 108-117). Ithaca, New York: Cornell University Press.

Fraenkel, J. R. & Wallen, N. E. (2000). *How to design and evaluate research in education* (4th ed.). Boston: McGraw-Hill.

Galili, I., Bendall, S. & Goldberg, F. (1993). The effect of prior knowledge and instruction on understanding image formation. *Journal of Research in Science Teaching,* **30**, 271-301.

Goldberg, F. M. & McDermott, L. C. (1986). Student difficulties in understanding image formation by a plane mirror. *The Physics Teacher*, **24**(8), 472-480.

Heller, P. & Finley, F. (1992). Variable uses of alternative conceptions: A case study in current electricity. *Journal of Research in Science Teaching*, **29**, 259-275.

Heywood, D. S. (2005). Primary trainee teachers' learning and teaching about light: Some pedagogical implications for initial teacher training. *International Journal of Science Education,* **27**(12), 1447-1475.

Iona, M. (1987). Why Johnny can't learn physics from textbooks I have known," Mario Iona's Acceptance speech for the 1986 Millikan Lecture Award presented by the American Association of Physics Teachers, Columbus, Ohio, 26 June 1986. *American Journal of Physics,* **55**, 299-307.

Ivowi, U. M. (1984). Misconception in physics amongst Nigerian secondary school students. *Physics Education*, **19**, 279-285.

Marshall, C. & Rossman, G. B. (1999). *Designing qualitative research* (3rd ed.). London: Sage Publications.

McDermott, L. C., Heron, P. R. L., Shaffer, P. S. & Stetzer, M. R. (2006). Improving the preparation of K-12 teachers through physics education research. American Journal of Physics, **74** (9), 763-767.

REASONING MODES, KNOWLEDGE ELEMENTS AND EPISTEMIC FRAMES IN PROBLEM SOLVING IN OPTICS

David Schuster
Western Michigan University

Adriana Undreiu
University of Virginia's College at Wise

Betty Adams
Western Michigan University

Abstract

We have investigated how students tackle physics problems in geometric optics involving ray construction, to try to understand the nature and origin of the surprisingly wide variety of their solution attempts. We find that students use various reasoning modes and knowledge elements in conjunction and their cognition may usefully be described as an interplay of principle-based reasoning and case-based reasoning. Their thinking draws on a mixture of basic principles, procedures, knowledge elements, specific cases and recalled result features, in a way that reflects certain epistemic frames of operating. Even though we usually present model solutions and teach problem solving as a systematic application of principles, this does not reflect the actual richness of either experts' or novices' cognition when tackling problems. Real cognition is much more complex. Associative thinking in terms of prior cases seems to be a strong natural tendency of both novices and experts. However, novices are not easily able to discriminate the specific from the general, and tend to lack epistemic awareness and metacognitive skills. Our research findings will be illustrated by examples and analysis of student thinking on a basic reflection problem. Implications for learning and instruction will be discussed..

Introduction

In physics courses students encounter both the general principles of a domain and specific example cases and worked problems. Thus students' overall knowledge gained from instruction will comprise both *general* and *specific* elements. It is only to be expected that both types will play roles in tackling subsequent problems. In addition, students will have their own general thinking tendencies and modes of operating, learned from everyday life and prior schooling, and may also bring experiential intuitions. In addition, instruction of variable quality may or may not have emphasized particular aspects of topic understanding. Given this background, what will happen when a particular student tackles a given physics problem? What reasoning modes and knowledge elements will be cued, in what situations, how, and why? What roles do general principles, procedures, specific cases, result features and intuitions play, and how do they interact? How does this compare to what experts do?

We undertook a cognitive study of problem solving in geometrical optics, investigating how students tackle reflection and refraction problems involving ray construction. Simpler problems involve just the basic phenomena of reflection or refraction, while more complex problems involve the image concept and location of images. The problems require direct application of basic laws in ray tracing, are of a conceptual and visual nature, and are essentially diagrammatic simulations of light behavior rather than being formula-based. Relatively little research has been done on student cognition in such optics problems. The area offers particularly rich possibilities for studying many facets of cognition.

An epistemic characteristic of science is to explain the physical world in terms of relatively few fundamental concepts and principles. We hope that students will also embrace this perspective in problem solving, applying basic relevant principles to construct solutions. We will refer to this approach as principle-based reasoning (PBR). It is reflected in our model solutions showing stepwise application of principles. While this may represent the physics structure of the solution, it does not reflect the actual richness of either experts' or novices' cognition when tackling problems. Hsu et. al. (2004) recently provided a useful general review of work on problem solving in physics; Bodner (2003) has noted how chemistry problem solutions likewise do not reflect actual thought processes.

Our study of problem solving in optics suggests that what instructors or experts may regard as involving a straightforward application of principles (laws of reflection and refraction) may not be what students see and do as they approach a problem. We find that many students instead try to 'map across' remembered results of earlier cases. We use the term case-base reasoning (CBR) to describe an approach where a problem solver recalls features of a similar case and tries to adapt them to the current problem. CBR can be useful and efficient for both novices and experts, but must be done with awareness of status and applicability conditions (Reif 1991, Kolodner 1993, Aamod 1994). A danger is that novices may simply map across surface features or specific case results, rather than transferring deep structure and procedures.

Research and Analysis

Intrigued by mystifying ray diagrams in student solutions and realizing they had unexpected difficulties, we undertook a phenomenographic [(Micari 2007) study of reasoning modes in problem solving in optics. We created sets of reflection and refraction problems to be done by ray tracing. Variants of a problem share a common underlying structure, while situational features and hence results differ. Students provided written solutions with explanations of their thinking path, difficulties, and confidence.

We probed students' reasoning on these problem sets in considerable depth, using a variety of methods, both qualitative and quantitative. Data sources comprised written problem-solving records, individual think-aloud interviews, clarification interviews, and researcher observations during class.

Table 1 gives questions considered in elucidating students' thinking. These aspects emerged as the study developed. We list them in two columns, reflecting reasoning based predominantly on principles or predominantly on cases. To analyze and interpret student cognition we developed a coding scheme, which emerged inductively from the study of many students' problem solving attempts plus interview data. Solutions and thinking were analyzed in 'episodes', and an individual's thinking path through a problem was represented as a sequence of such coded cognitive elements.

Findings

The study shows that students bring in a great variety of knowledge elements and reasoning modes, at different stages of the problem, for different purposes, and with different status. Despite this complexity and variation, it became evident that basic principles and remembered cases both play major roles in solution attempts. Thus students' thinking may be usefully divided into two main categories: principle-based reasoning (PBR) and case-based reasoning (CBR). Rarely if ever do people use only one of these in its pure form; for novices and experts alike there is usually an interplay of both. An interesting aspect is that students may alternate between PBR and CBR as they work on a problem, or solve part of it one way and part the other, often without being aware of this. Overall, students' epistemological stances toward science and the problem task seemed to shape what they did and why. There seemed to work in tacit 'epistemic frames', usually either a principle-based frame or a cased-based frame. Within these they played particular 'epistemic games' (Tuminaro and Redish, 2007) using a variety of 'pieces', 'moves' and 'rules', and sometimes switched games. This epistemic perspective was a useful way for us to make sense of the cognitive variety we encountered.

Table 1. Analysis Considerations

These questions served to guide our analysis and interpretation of student solutions and thinking.

Regarding principles and their use	Regarding cases and their use
• Are concepts & principles/procedures drawn upon? If so, which ones? • Are they appropriately related to the problem or not? Why? • Declarative: are concepts & principles stated correctly? If incorrectly, how and why? • Procedural: Principles applied correctly? If not, what errors are made and why? • What role do principles serve in this attempt? To produce the answer? To evaluate a case-based result? To justify a result? • Epistemically aware of role of principles in science?	• Are related cases or previous problems drawn upon? If so, which? • How do they relate to the new problem? Similarities & differences? • What *type* of case knowledge is invoked? Surface features? Result features? Solution method? Other? • What *role* does a previous case result serve? Predict expected result? Provide result directly? A basis for result adjustment? Check result obtained from principles? • How is case knowledge used? Mapped 'as is' over to the new? Completely or partially? Mapped but adjusted? • What role does a recalled case *solution method* serve? Procedure to imitate? Guide to implementation of basic principles? • Aware of status and applicability conditions of case features?
Principle-based and case-based aspects: comparison and resolution	
• Are both principle-based and case-based aspects present in an attempt? Do they appear to conflict or agree? • Do students show awareness of comparison or not? Why? • How does apparent agreement or conflict affect students' attempts, as well as their confidence in their work? • If aspects are in conflict, do students try to resolve it? If so, how? Do they try to adjust something to fit? What? • If they cannot resolve a conflict, what do they do? Which mode is 'trusted' more i.e. 'wins', and why?	
Confidence level with reasons	
• How confident is a student about their own solution attempt (be it correct or not) and why?	

Example: cognition in a simple reflection problem

Even for apparently simple problems thinking can be amazingly varied and interesting, and students have unexpected difficulties. Therefore, to make this point in a short paper, we choose one of the simplest problem variants as an example. (We have data on many more problems, on reflection, refraction and images).

Students had learned the law of reflection, practiced it in single-reflection problems, then done the classic 90° corner-mirror problem involving two successive reflections (Fig.1.a.), followed by discussion of the solution.

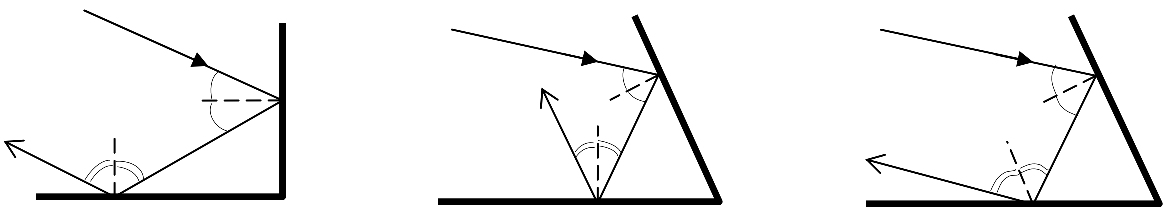


FIGURE 1. **a**. Case encountered in instruction . **b.** Variant used in examination. c. A student solution using both PBR and CBR.

For the next test, students were given a variant with a different angle between mirrors. To trace the ray path, one needs to conceptualize the task as applying the law of reflection twice in succession (Fig.1.b). What kinds of ray solutions do students construct and what modes of reasoning lie behind them? We will not describe the whole spectrum here, but illustrate some interesting aspects of student reasoning.

As expected for a simple problem, a high proportion (three quarters) of students ended up with a correct diagram. Even where students described their final work in terms of applying principles, interviews and written reflections indicated that they rarely drew on principles alone. Cued by similarities with the corner mirror case, students recalled specific features which could be either beneficial or not, depending on whether they also noticed the differences. A variety of reasoning scenarios emerged from our data, and are discussed below as primarily principle-based, case-based, or an interplay, illustrated with interview extracts.

I. PBR as primary approach

In many cases PBR was the primary approach, with case knowledge playing various roles.

• Principled solution, with previous case result as expectation or evaluation

“ *My .. thought was that light will come out parallel in the opposite direction.”* An expectation from a previous special case. The recall of case features may confuse some students but help others. Some students experience an ‘aha!’ moment. They compare cases, and recognize both similarities and differences (Marton, 2006). They can then accept a principled but unfamiliar result with confidence. *“ I checked my work and I told to myself that’s] ok because the situations are different.”* However, conflicts often remain unresolved by students. In such situations, even with a correct PBR solution (Fig.1.b), students lose confidence in their work. *“I thought the lines had to be parallel…in the previous problem… … this is what I was picturing… my drawing it ended up being a triangle, they weren’t parallel, so I second-guessed myself.”* She did not think about case differences.

• Principles/procedures recalled via drawing on previous case procedures

“First I thought of how we did the wall reflection problems in class.” The similarity and familiarity seems beneficial here, helping a student to implement the desired solution method. *“ ..but this was different because the mirrors were not at right angles”.* Thus she also recognizes the difference between cases:.” Thus she can account for the different result features, and has confidence in her principled solution, despite the fact that the result is unfamiliar.

II. CBR as primary approach

For other students cases were dominant in their thinking, and principles played a secondary role.

• *Result imitation.* Based on recognized similarities, students interpreted a new problem as a familiar–looking case that must therefore have a familiar result, rather than constructing a solution from principles. *“I remembered seeing it parallel so I tried to draw them parallel…I sure remembered seeing it, I could visualize it on the board.”* They thus ‘map across’ results (fig. 1c).

• *Case result adaptation.* When they do recognize case differences students may guess at how to accommodate that difference. *“ .. this mirror is slanted so it made this one to head a little bit differently, so they are not exactly parallel as they are here”.* Rather than using principles students simply adjust the angle of the outgoing ray so that it would no longer look parallel. Adaptation is preferable to imitation, but the absence of epistemic and metacognitive awareness is of concern.

III. Interplay of PBR and CBR

Another characteristic of the student diagram in fig.1c. is that it shows a combination of PBR and CBR.

The student uses PBR for the first reflection, but for the second she then uses CBR, reproducing the outgoing ray from the previous case. She also then tries to justify her result by backward application of principles, which forces her to draw an incorrect normal to allow the result she expected.

Conclusions and Implications for Instruction

There is a fascinating complexity to thinking and problem-solving. Behind the ray diagrams offered as solutions ‘hides’ a combination of reasoning modes and strategies based on principles, cases, intuitive ideas and everyday thinking strategies. Solving problems using compiled knowledge from similar cases (CBR) seems to be a strong tendency of both novices and experts (Reif 1991). However there is a difference between the way experts and novices use case knowledge. Used with awareness of applicability conditions, case differences, and generality or specificity, CBR can be a powerful and efficient strategy (Kolodner 1993, Aamod 1994) in both science and everyday life.

Because cognition has so many complex facets the task facing the student is greater than that addressed by prevailing instruction, which focuses mostly on content, not cognition. This tendency may stem partly from an inappropriate epistemological framework of 'what it is all about'. From prior schooling they may have come to view learning as memorization, exams as involving factual recall and 'seen' problems, and answers as more important than understanding method. Implications for more effective instruction are that it should involve not just content but also cognition and epistemology. It should incorporate what we currently know about thinking and the nature of expertise, and promote reflection, metacognition, and epistemic awareness. This includes teaching about both PBR and CBR, in context, and making such thinking 'visible' (Collins et. al. 1991).

References

Aamodt, A., & Plaza, E. (1994). Case-based reasoning: Foundational issues, methodological variations, and system approaches. *Artificial Intelligence Communications, 7*(1), 39-59.

Bodner, G. (2003). Problem solving: The difference between what we do and what we tell students to do. *U. Chem. Ed., 7*, 37-45.

Collins. A, Seely Brown, J. and Holum, A (1991). *Cognitive Apprenticeship: Making Thinking Visible.* American Educator: The Professional Journal of the American Federation of Teachers, v15 n3 6-11, 38-46.

Kolodner, J. (1993). *Case-based reasoning.* San Mateo, CA: Morgan Kaufmann.

Hsu, L., Brewer, E., Foster, T. M., & Harper, K. A. (2004). Resource letter RPS-1: Research in problem solving. *American Journal of Physics, 72*(9), 1147-1156.

Marton, F. (2006). Sameness and difference in transfer. *The Journal of the Learning Sciences, 15*(4), 499-535.

Micari, M., Light, G., Calkins, S., & Streitwieser, B. (2007). Assessment beyond performance: Phenomenography in educational evaluation. *American Journal of Evaluation, 28*, 458-476.

Reif, F., & Larkin, J. (1991). Cognition in scientific and everyday domains: Comparison and learning implications. *Journal of Research in Science Teaching, 28*(9), 733-760.

Tuminaro, J., & Redish, E. (2007). Elements of a cognitive model of physics problem solving: Epistemic games. *Physical Review Special Topics - Physics Education Research, 3*(2), 020101.

Allan Collins, John Seely Brown, and Ann Holum (1991) Cognitive Apprenticeship: Making Thinking Visible. *American Educator: The Professional Journal of the American Federation of Teachers, 15* (3) 6-11, 38-46.

HOW TO COPE WITH GAUSS'S ERRORS? LEARNERS' OBSTACLES AND THEIR POTENTIAL ROOTS IN THE HISTORY OF DATA TREATMENT

Susanne Heinicke
University of Oldenburg, Germany

Falk Riess
University of Oldenburg, Germany

Abstract

An online questionnaire on the understanding of measurement and experimental practice was distributed to 160 students from ten different German universities prior to first laboratory instruction. In accordance with prior studies the data show a lack of deeper understanding of the routines and terminology involved when handling measurement data.

Although we know a certain amount of the students' difficulties on the handling of measurement data und uncertainty, however, little do we know as to the reasons, why. We therefore decided to both remodel the students' concepts of measurement and uncertainty as well as to analyse the traditional statistical (Gaussian) formalism of error calculation for its underlying concepts and compare these two for similarities and opposites.

We believe the latter is inconsistent with the students' concepts and might constitute obstacles in the students' learning process. We want to discuss how information on the historical background of the traditional approach can support our understanding of the students' notions. Further we wish to raise the question if the historical background can be utilized in the teaching of data treatment to support the students' understanding or whether pedagogically and scientifically justifiable alternatives to the traditional approach should better be introduced in the laboratory context.

Introduction

Over the past 15 years a number of studies have been carried out internationally on the students' understanding of scientific measurement und measurement uncertainty. A kind of starting shot to this research was given by a study of Séré et al. in 1993 among 20 first year students who took a special class on statistics. They summarize:

> "Their [the students'] analysis shows that the students establish a hierarchy in the series of measurements and do not really understand the need to make several measurements. ... Random and systematic errors are not distinguished. The general view is that, the more measurements one makes, the `better´ the result is, without understanding the nature of this `better´."

From subsequent studies as well as our own we learned that there is in fact an even larger variety of difficulties that learners face when it comes to the understanding of measurement and uncertainty. However, having some insight into the kind of the students' struggles, little do we know as to the reasons *why* they struggle.

There are surely many ways to obtain a deeper understanding of the learners' difficulties to understand the nature of measurement uncertainty. The approach chosen here is to analyze the very matter that we are trying to teach to the students: the handling of measurement uncertainty. This concerns both the scientific content of the methods as well as the historical context in which the methods were first derived. Especially the historical background has not yet been investigated by previous studies. Yet, we believe that this analysis holds valuable information in two ways: firstly in order to understand how the methods, technical terms and category systems that we use as a matter of course emerged and what tacit assumptions might implicitly underlie these methods (Hon 2005). Secondly, following Piaget's genetic epistemology, the knowledge of the experts' understandings can again inform our understanding of the learner perspectives as the students' notions often reflect the examples of early expert scientists.

In the following the term "traditional" error analysis refers to the approach to data handling developed by Gauss and others at the beginning of the 19th century. The approach is based on the statistical treatment of measurement data and indicates the probability that a certain result will occur if the reading is taken an infinite number of times. Fundamental aspects of the approach are (1) a large number of data (in the mathematical model $N \rightarrow \infty$), (2) the modeling of data using the normal distribution function, (3) the computation of the final result in terms of arithmetic mean (best estimate) and standard uncertainty (measure of variability of data) and (4) the calculation of the overall random uncertainty in terms of error propagation.

Rationale

In the following we will compare the learners' notions and difficulties that were assessed in previous studies and our own to the very matter that we are trying to get across in the error analysis taking into account both the scientific content and the historical perspective. Following from this comparison we will draw conclusions on possible implications on Science Teaching with special regard to laboratory instruction. As the learner's perspective is complex and manifold we chose three examples as illustration that were also tackled by the résumé of Séré et al.:

1. The Terminology commonly used in textbooks and by the learners as fundamental components of the learning process and building of concepts.
2. The different reasons of repeating a measurement as fundamental aspects of the traditional statistical approach to data treatment.
3. The mere ability to apply the routines of error calculation presented in the laboratory compared to the development of a deeper understanding of the routines and the nature of measurement data.

Methods

To attain a picture of the situation of German students at school and university we developed a questionnaire based on the instrument as described in Allie et al. (1998). In the case of university students we distributed the questionnaire online to 160 students prior and 110 students post to first year laboratory instruction at ten different German universities. The questionnaire contained questions on the collection, analysis, interpretation and comparison of data in a given laboratory context, on different aspects of the Nature of Science and on the terminology commonly involved. Based on the students' answers a coding scheme was derived. The analysis of the learners' perspective in comparison to the analysis of the scientific perspective form the "bottom line" of the triangle of the educational reconstruction. From the comparison of the two possible implications on science teaching and laboratory instructions are drawn.

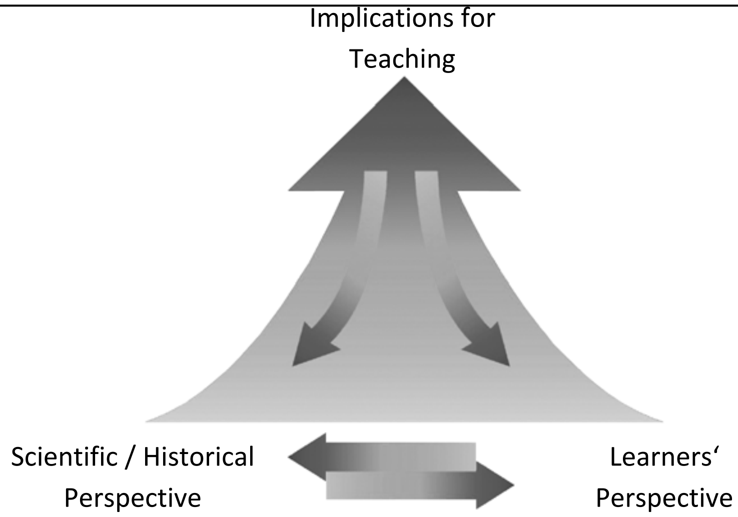


Figure 1: Model of the Educational Reconstruction (Duit et al. 2005)

For the analysis of the scientific content and historical background of the traditional methods we used both original texts especially from 1750-1850 (the period in which the Method of Least Squares was developed), school and university textbooks as well as international standards, especially the Guide to the expression of Uncertainty in Measurement (GUM) and the Vocabulaire International de Métrologie (VIM).

Results

Terminology

Following Séré et al. (1993) and a study of Tomlinson et al. (2001) on the language of error, we analyzed the terminology (1) used in common textbooks and (2) used by students themselves.

For the latter case we analyzed what terms the students used when defining the spread of a measurement series or when naming the "+/- value commonly added to measured quantities". The students were also asked to express the meaning of the most common terms *measurement error*, *inaccuracy* and *uncertainty*. The analysis showed three interesting results:

1. The students use a large variety of more than 20 different terms both in the pre and in the posttest, e.g.: *measurement error, uncertainty, deviation, Standard uncertainty, inaccuracy, sigma, delta, Gaussian error, Student factor, systematic / random error* etc..
2. The term most often used in the pretest was the term „measurement error" whereas in the posttest „measurement uncertainty" was used slightly more frequently.
3. The meanings the students attributed to the terms differed completely from term to term.
    - E.g. for 65% of the students the term *measurement error* denotes some erroneous proceeding on behalf of the experimenter like a mistake in the measuring process, the reading off or the set up of the experiment. 15% defined the term as the deviation of the measured from the true value.
    - The term *measurement (in)accuracy* was mostly related to the limited sensitivity or accuracy of a measuring device (75%).

- Measurement uncertainty was the term least well-known to the students. Its definitions given by the students showed a large variety. About 25% of the students stated that the term meant the experimenters were uncertain about what to do in the experiment, e.g. what device to use, what method to choose and which reading to take. Others stated that the term denoted an interval, the result of the error calculation or the random error, each 10% respectively.

Referring to the first two aspects the findings on the learners' perspective were compared to the situation of common textbooks. Firstly we found a similarity in the variety of different terms in the analyzed textbooks. Secondly these different terms were most often used synonymously within the same textbook. A typical laboratory manual (2009) for example reads:

> „If in the following we use the word *error*, we do not mean that something has been done wrong. Rather, the term *error* denotes the uncertainty of a measurement value. To avoid linguistic confusion, it is suggested [i.e. by intern. standards)]… to speak of *deviations* rather than of *errors* …In the daily routine of physics however, this has not yet been established. Therefore we will stick to the term *error* in the sense of the word as discussed above".

A very recent textbook (2008, emphasis added) states:

> "Here [when measuring] … *measurement uncertainties* remain. They are denoted as *measurement errors*. … If measurements are repeated in a series deviations occur. The measurement values are scattered around an average value. Such *measurement inaccuracies* are called *statistical errors*."

Comparing these results to the learners' perspective we see a tension between the synonymous usage of terms in school- and textbooks and the students' discriminative definitions to the relevant terms. Taking into account the very different associations the learners have when defining the terms a synonymous usage seems neither helpful to them nor for the complex matter that is to be discussed in the texts. Considering the history of the common terminology of error treatment we find that the term error can easily be traced back way in time. Galileo for example writes in his *Dialogue Concerning the Two Chief World System* (1632):

> **„**If now the astronomers‘ observations were true and the author‘s calculations without error, then in all cases it would have to follow that the particular [correct] distance results.“

Gauss likewise uses the term *error*, e.g. in his treatise on his approach to error calculation (1809 and 1821). In his Theoria Combinationis (1821) he also once uses the term *uncertainty* referring to the *uncertainty of human vision* (cf. to learners' association of uncertainty of experimenter). The meaning of the term error is historically rooted in the scientific concept that to every measurement there is a true value available which is theoretically accessible and allocable if all disruptive influencing, erroneous factors are eliminated. Such an approach blends in well with the traditional laboratory of recipe-like instruction where the answers to experimental exercises are already known in terms of authoritative values which can be looked up in the manual or textbook. A subsequent question would be if this 'true value' does in any case actually exist and if such a concept is in fact helpful to the students when developing an understanding of assessing the uncertainty of a measurement result from its experimental determination rather than assessing its measurement error by extern values. The terminological and conceptual separation between the *error* and *uncertainty* used here in contrast to the synonymous usage and lacking separation accuracy in textbooks and laboratory practice was defined in the *Vocabulaire International des Mésures (VIM)* and in the *Guide to the Expression of Uncertainty in Measurement (GUM).* These international standard suggest distinct terminological definitions. In Note 2 the GUM (2008) states:

> "In this Guide, great care is taken to distinguish between the terms '*error*' and *'uncertainty'*. They are not synonyms, but represent completely different concepts; they should not be confused with one another or misused."

The two terms are defined as follows:

- The term *measurement error* refers to a measured quantity value minus a reference quantity value. As the reference value is most often unknown and unknowable such is the quantity of the measurement error.

- The term *measurement uncertainty* on the other hand denotes a parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand.

To conclude about the terminology: For the learners the different terms denote different concepts which find some of their aspects reflected in the definitions given by the GUM. From the learners' perspective, the interchangeable usage does not appear to be helpful as different notions underlie them. The claim of the GUM to use them as completely different concepts however, has not yet been implemented in textbooks and lab manuals where the terms are still used interchangably and thus add to the confusion of the students. As a trend in many textbooks in the last 30 years a change in terminology from *error* to *uncertainty* occured. However, this change implemented in the GUM refers not only to a terminological but a conceptual one.

Reasons for repetition

A fundamental question concerning the accuracy and precision of a measurement result is now, how measurement errors or measurement uncertainties can be reduced, how the accuracy and precision of a measurement result has to be assessed and for what purpose a reading should be repeated. There are different points of view one can advance here. The approach that Gauss favored is rooted in the understanding that a larger sample of readings will constitute a more precise result when taking its arithmetic mean and thus the repetition of readings form the baseline of the traditional statistical approach to error treatment.

Scientific-Historical Perspective

However, this position was not so straight forward for his time (and as our empirical data will show neither is for ours). To take just another example, Coulomb published his results on the Laws of Electro- and Magnetostatics in a paper in 1785. To establish his idea of an inverse square relationship between distance and force he presents his results as follows (Coulomb, 1785): „I here present only some assays which are easy to repeat and which at once reveal the Law of repulsion.“ He subsequently gives three data pairs which are all in accordance with his theory. The question arises how Coulomb could present such a small number of readings to establish such a fundamental law of physics and if this procedure was acceptable for his time. There are several answers for this. We want to highlight here only one which has to do with the understanding of the accuracy of experimental results. Taking again Galileo‘s quote that the necessary condition for an exact result is a true observation and a faultless computation of results, thus, the reliability of a measurement does not manifest itself in a large number of data but in its closeness to the true value, its *accuracy*. For the time and situation of Coulomb one could moreover say that he would rather have undermined his authority as an excellent experimenter had he given more data than he actually did (Heering 2006).

One could now infer, that in fact Coulomb probably took more readings than he subsequentely published. This might be indicated by his comment: „I here present only *some* assays…“ He might have taken some more readings first to practice, or a couple of values from which he then chose those more consistent with theory.

To conclude, Coulomb's understanding of the purpose of repeating a measurement could be that a repetition is only another check-up of the result(s) already obtained, a matter of training before taking the "real" reading or a method to equip oneself with a set of numerical options to chose from according to e.g. their agreement with his theoretical expectations.

Gauss on the other hand endorsed Coulomb's view about the check-up role of a repeated reading. However, he limits this role to the case of flawless measurements as he writes in his first essay on the Method of Least Squares (1809): "If the astronomical observations … were absolutely correct, the elements also … would be strictly accurate, … therefore, if other observations were used [added], they might be confirmed, but not corrected." In this case, as Gauss pointed out, additional experimental data would not add any quantitative information about the value of the measurand but only help to confirm or reject the result already obtained. "*But*", he continued – and this is so to say the turning point of the story – „*since all our measurements and observations* ***are nothing more than approximations*** *to the truth, …. The highest aim … must be to approximate, as nearly as possible, to the truth. But this can be accomplished in no other way than by a* ***suitable combination of more observations*** *than the number absolutely requisite for the determination of the unknown quantities* [emphasis added]". In this understanding, which differs distinctively from Coulomb's and many of Gauss's fellow scientists', consequently, the collection of more data than necessary regarding the mathematics of the problem now adds to the quantitative information about the value of the measurand. This alternative understanding is on the one hand rooted in the understanding that all available data adds to the knowledge of the final result, an understanding that Allie et al. (1998) denote with the term "set paradigm", whereas from Coulomb's point of view each data point is treated independently from each other, denoted by Allie et al. (1998) with the term "point paradigm". On the other hand Gauss's understanding is based on the idea of a randomness of data distribution. Only if the deviations from the target value are randomly distributed the procedure of taking the mean of a large sample of readings will lead to a more precise result than an average single reading. Without this postulated randomness of data distribution, Coulomb on the other hand rather assesses his data in terms of their accuracy, i.e. their agreement with the expected result. In this understanding taking the mean of a measurement series would not make much sense exceeding the mere check-up of results. In fact some single readings will always be more accurate, i.e. closer to the true value, than the mean of the whole series. The difficulty that Coulomb pursuits is merely to find out to which readings this corresponds.

Necessary for this new approach towards measurement precision and the repetition of readings is however, a mathematical tool to handle an overdetermined data set. Gauss and others present this tool in terms of the method of least squares which forms the basis of the traditional statistical error calculus.

In summary, one of the fundamental concepts that underlie the traditional statistical error calculus is the notion that the mean of a larger sample of data will yield a more precise result than a single reading could provide. This only makes sense given the randomness of data distribution. Otherwise other arguments could be employed how to assess the accuracy or precision of results or how to single out the most adequate values.

Learners' Perspective

After analyzing the historical background and fundamental concepts that underlie the scientific content, we compare these results to the perspective of the learners. In a study of Allie, Buffler & Lubben (1998) on more than 120 first year students at the University of Cape Town, the students were asked whether and how often a certain measurement should be repeated and to justify their choices. Allie et al. (1998) analyzed the following arguments in the students' answers:

Question: Should the students repeat their measurement?

- Once is enough, they got their result.
- They should take a second reading to check the result.
- They should take a couple of times to practice and then take the reading.
- They should take a couple of readings to be sure they got the right value.
- They should take a couple of readings to avoid / detect errors.
- They should take a couple of readings to reduce systematic / random errors.

In accordance with the results of Allie et al. (1998) we found in our own study the following main arguments (numbers rounded off to indicate uncertainty of interpretation, multiple answers possible):

- (30%) Whether one should make a repetition depends on the result's **deviation** from the target value.
- (15%) When repeating results can be **checked.**
- (40%) A repetition **increases accuracy / precision** [German both: Genauigkeit].
- (15%) When repeating **possible errors** can be detected.
- <15%: When repeating one can account for statistical **(random) fluctuations.**

Here, the first two arguments closely compare to the understanding that we drew of Coulomb's work. The underlying notion is that the assessment of accuracy is based on the comparison of numerical values only. Traditional laboratories where students follow a very structured, recipe-like instruction in order to determine a value known to them or at least to the instructor a priori, support or at least do not challenge such an understanding. The number of repetitions to be taken is usually given (commonly 3, 5, 6 or 10 readings) without explanation to the context. More than half of the students argue that repetition increases accuracy or precision or helps to detect errors without giving further explanation how this passes off. This argument is in accordance with the results of Séré et al. (1993) as quoted above that "the general view is that, the more measurements one makes, the `better´ the result is, without understanding the nature of this `better´". Only less than 15% of the students in our study established a relationship in their answers between the repetition of readings and the randomness of fluctuations in data sets, the argument that turned out to be a fundamental aspect of Gauss's approach to error treatment and to the formulation of the statistical approach of the traditional error calculus.

Application versus understanding

The preceding analysis of different obstacles in the learners' way to develop an understanding of the handling of measurement uncertainty raises the question to what degree the learners actually are enabled to develop a deeper understanding of the routines of error treatment that are taught in the laboratories or to what degree they merely learn to apply the routines by rules of thumb or by rote.

At this stage of the analysis of our data the answer to this question cannot satisfactorily be given. However, as an indication we chose the students' answers to a question on the comparison of the results of two different groups of students in the laboratory:

Question: Do the results of the two groups (101.120 C/mol and 102.034,6792 C/mol, relative uncertainty ~2%) agree?

	pretest	posttest
– **Comparison of numerical values („Values agree / are too far apart.“).**	**51%**	**17%**
– Rules of Thumb ("5% deviation is still close enough.").	10%	13%
– **Refer to error calculation ("This can only be decided upon after error calculation.")**	**15%**	**37%**
– Comparison of resulting intervals ("Error / uncertainty / data intervals overlap."):	19%	18%
– Quantitative estimation ("Deviations are in dimension of uncertainties of single readings."):	2%	5%

Interestingly, in three of the five groups the overall percentage of answers remains relatively constant from pre- to posttest. In the two groups highlighted, however, there is a clear decrease and increase in the proportions. On the one hand the mere numerical comparisons decrease from more than half of the answers to less than a fifth. On the other hand we see a clear increase in answers advancing the view that nothing can be said or estimated about the possible agreement of the two values before any calculation is done on their errors / uncertainties. It is important to note here that prior to this question all detailed information about the measuring process had been given to the students in the study.

This result suggests that in accordance with findings of Séré e tal. (1993) the students seem to acquire the presented routines of traditional error treatment as a rule of computation that has to be carried out before any inference about the results can be made. Estimation skills are not addressed in this approach and because of the complexity of the computation the approach seems to be not transparent to the learners. In an interview with one of the students in the study she stated: "At the end the result of error calculation just appears. Well, I couldn't estimate it beforehand. This is such a complicated formula. And to estimate the individual measurement errors is so arbitrary. I often feel like I just figure out something to make it fit. It is also that I already know what has to come out of it in the end. … And if it still doesn't fit in the end I have to fudge a reason to explain the deviation."

Conclusions and Implications

Historically the discovery of randomness in data distribution gave rise to the development of a statistical error calculus by Gauss and others. The calculus was based on the probability and gambling theory and on the fundamental concept that a repetition in measurement reduces random fluctuations and thus produces a more precise best estimate. Our study as well as studies of Séré et al. (1993 and Allie et al. (1998) have shown however, that students lack an appreciation of this basic concepts of randomness and the purpose of the repetition of readings, confuse random and systematical errors and after instruction rely on carrying out the computations involved in the error calculus without developing a deeper understanding of the routines or an ability of qualitatively and quantitatively estimating the precision of measurement results.

We therefore suggest to explicitly discuss the implicit concepts and ideas that underlie the traditional approach (e.g. the role of recovering randomness in data distribution) to data treatment in classes and laboratories Another possible step that will follow our study is to implement the approach as described in the Guide to the Expression of Measurement Uncertainty to include both random as well as systematical deviations into the computation and use the well defined terminology as suggested in the VIM (see also Buffler et al. 2008) and to compare those implementations to the traditional approach common in most university laboratories.

References

Allie, S., Buffler, A., Campbell, B., Lubben, F. (1998). First year physics students' perceptions of the quality of experimental measurements. International Journal of Science Education, 20 (4), 447-459.

Börsch, A. & Simon, P. (1887). Abhandlungen zur Methode der kleinsten Quadrate von Carl Friedrich Gauss. Berlin : P.Stankiewicz, 1-27 (Theoria Combinationis 1821), 92-117 (Theoria Motus 1809).

Buffler, A., Allie, S. & Lubben, F. (2008). Teaching Measurement and Uncertainty the GUM way. The Physics Teacher, 46 (12), 539-543.

Coulomb, C. (1785). Premier Mémoire sur l'Electricité et le Magnétisme. Histoire de l'Académie Royale des Sciences, 569-577.

Coulomb, C. (1785). Deuxième Mémoire sur l'Electricité et le Magnétisme. Histoire de l'Académie Royale des Sciences, 578-611.

Duit, R, Gropengiesser, H. & Kattmann, U. (2005). Towards science education research that is relevant for improving practice: The model of educational reconstruction. In: Developing standards in Research on Science Education – Fischer (ed). London: Taylor & Francis Group.

Galilei, G. (1632). Dialogue Concerning the Two Chief World Systems. Translated by Drake, S., University of California Press, 1953.

Heering, P. (2006). Regular Twists: Replicating Coulomb's Wire-Torsion Experiments. In: Physics in Perspective 8, 52-63.

Hon, G. (2005). Putting error to (historical) work: error as a tell-tale in the studies of Kepler and Galileo. Centaurus 46, 58-81.

International Organization for Standardization (2008). Guide to the Expression of Uncertainty in Measurement (GUM). Switzerland.

Séré, M.-G., Journeaux, R. & Larcher, C. (1993). Learning the statistical analysis of measurement error. International Journal of Science Education, 15 (4), 427-438.

Tomlinson, J., Dyson, P. & Garratt, J. (2001). Student misconceptions of the language of error. University Chemistry Education, 5 (1), 1-8.

TURKISH STUDENTS' FORCE MEANINGS

Muhsin Menekse
Arizona State University

Douglas B. Clark
Vanderbilt University

Gökhan Özdemir
Niğde University

Cynthia Dangelo
Arizona State University

Sharon Scheligh
East Carolina University

Abstract

What are Turkish pre, elementary, middle, and high school students' force ideas? And, how do Turkish students' non-normative force ideas differ or be similar to the well-known force misconceptions reported in the literature? Students have false and persistent beliefs about the physical world and they struggle with challenging misconceptions based on their perceptions of everyday experiences. The current study applied the coding schemes from diSessa, Gillespie, and Esterly (2004) to interviews conducted with 78 students from two cities in Turkey in order to investigate students' understandings of force concept. Specifically, this study focuses on Turkish students' force meanings in different situations and the commonalities of the ideas at various age groups. The results show that there are significant differences, in terms of force meanings, among grade levels and high school tracks but no difference by gender nor city students live. Also, the distribution of force meanings across grade levels indicate remarkable findings, such as force meanings are uniformly distributed for both middle and high school students. Overall, this study, with its large sample size and in-depth interview questions, offers a significant contribution to the limited literature of Turkish students' force-meanings.

Background and Purpose

The force concept is one of the most important and complex topics in physics. All physics curricula, at various grade levels, require the normative understanding of this fundamental concept in order to study advanced concepts (such as linear momentum or rotational dynamics). However, many highly cited studies show that force-related concepts are not well understood by the majority of students (e.g., Chi, 2005; Clement, 1982; Cooke & Breedin, 1994; diSessa, Gillespie and Esterly (henceforth referred to as DG&E), 2004; Halloun & Hestenes, 1985; Ioannides & Vosniadou (henceforth referred to as I&V), 2002; McDermott, 1984; Watts 1983). Students have false and persistent beliefs about the physical world and they struggle with challenging misconceptions based on their perceptions of everyday experiences. For example, many students demonstrate the misconception that an increasing force is required in order to accelerate objects horizontally (e.g., Watts & Zylbersztajn, 1981). Also, as another example, students think a heavier object exerts greater force on the lighter one in an interaction because it dominantly affects the lighter's opposition (e.g., Gunstone & Watts, 1985). Students acquire these naive conceptions from the physical world and these ideas are strengthened by everyday experiences and actions like throwing, lifting, pulling, or pushing and observing objects in action. These observations and experiences strongly shape students' perceptions and understanding about scientific knowledge, and causing difficulties to develop normative concepts (Anderson, 2007).

The current study uses the set of ten interview questions, which DG&E condensed and modified from I&V for their quasi-replication research, in order to investigate Turkish students' understandings of force. The main purpose of this study is to document Turkish students' force ideas and to compare the findings with core alternative frameworks stated frequently in the literature. Specifically, this study focuses on students' force ideas in different situations and the commonalities of the ideas at various age groups. The two main research questions of this study therefore investigate: (1) what are Turkish pre, elementary, middle, and high school students' force ideas? (2) How are Turkish students' non-normative force ideas similar and/or different from the misconceptions observed in studies of English-speaking students?

Rationale

Research show students have or develop many force misconceptions based on their everyday experiences and intuitive ideas. One can classify these misconceptions in different categories, such as ideas specific to certain age groups, or ideas ordered from most resistant to easy to change. Duit's (2007) excellent bibliography documents hundreds of research studies regarding misconceptions and conceptual change. Unfortunately, however, the majority of these studies have been conducted with English speaking students and there is limited number of studies with Turkish students regarding force concept. Thus, it is difficult to draw strong conclusions about Turkish students' force meanings with the existing literature. Also, the differences in age-groups and research methods among the studies with Turkish kids make it difficult to generalize.

In this study, we interviewed Turkish students of various ages, with a large sample size (N = 78), to investigate their ideas related to force concept. We hypothesized that because culture and language related factors might potentially influence students' interpretation of force, interview analysis may indicate diverse results for Turkish students. They may express some force ideas different (or same) than those previously documented in the literature.

Methods

Participants

This study involves 78 students from two cities in Turkey, 32 from Ankara and 46 from Gaziantep. The students were from four different grade levels including: 8 pre-school students per city, 8 elementary students per city, 8 middle students per city, and 8 high school students from Ankara and 23 from Gaziantep. The mean student ages were 5, 10, 13, and 16 years, respectively. Half of the students were girls and half were boys. The interviewers were two Turkish native speakers, and doctoral students. No more than three students at any age group were selected from the same school. All students were interviewed individually for about 20-25 minutes. All interviews were videotaped, transcribed, and translated into English.

Instrument

This study uses the set of ten questions that DG&E condensed and modified from I&V for their quasi-replication research (see figure 1). Each set involves three questions: two simple yes/no questions and one comparison question. The simple questions directly asking for whether there is force on a specified stone or not. The comparison questions consist of two drawings comparing the different situations to investigate the contexts in which the students would refer to force/s and how they would explain those force/s.

In each set, the typical question was, "Is there a force on this stone? Why?" and the specific drawing was shown while asking the question. After asking the same questions for the second drawing in each set, students were asked, "Is the force on this stone (in the first picture) the same or different than the force on this stone (in the second picture)? Why?" Such comparison questions were asked as long as students say "yes" for both of the simple questions. The comparison question, when applicable, provides more information related to the student's understanding of force in terms of strength and contextual-related differences.

Analysis

Interviews were analyzed by using DG&E coding scheme, which was adapted from I&V's coding schema. DG&E' schema does "not include consideration of explanations, as I&V did" (p.867) in order to avoid possible bias in coding. Instead, they used more "holistic model mapping" technique to code interviews. According to DG&E's schema students' responses to each set of questions assigned one of the seven force meanings categories: (1) internal force (2) internal force affected by movement (3) internal and acquired (4) acquired (5) acquired and force of push-pull (6) force of push-pull (7) gravity and others. Table 1 represents a sample coding schema for question set 1. Each question set coded by comparing of amount of force on each stone with possible caveats and all possible matches calculated for each the seven force meanings (see Özdemir & Clark 2009 for more details in coding). Each interview was coded individually by two different coders (inter-rater reliability was approximately 93%) and then any differences were discussed. The final code was agreed upon for each student for each question set.

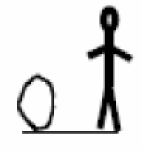
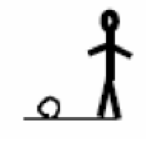
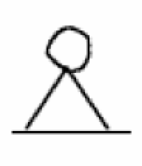
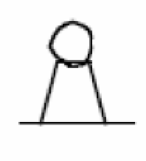
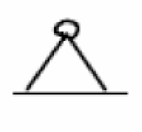
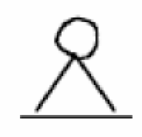
Question Set	Drawing A	Question A	Drawing B	Question B	Comparison Question
1		"This stone is standing on the ground. Is there a force on this stone? Why?"		"This stone is standing on the ground. Is there a force on this stone? Why?"	"Is the force on this stone the same or different than the force on this stone Why"
2		"This stone is standing on a hill. It is unstable. That means it could easily fall down. Is there a force on the stone? Why?"		"This stone is standing on a hill. It is stable. That means it won't easily fall down. Is there a force on the stone? Why?"	"Is the force on this stone the same or different than the force on this stone? Why"
3		"This stone is standing on a hill. It is unstable. That means it could easily fall down. Is there a force on the stone? Why?"		"This stone is standing on a hill. It is unstable. That means it could easily fall down. Is there a force on the stone? Why?"	"Is the force on this stone the same or different than the force on this stone? Why"

Figure 1. Question Sets 1, 2 & 3 (total of 10) combined by DG&E based on I&V's questions.

Results

The preliminary results show the variation in force meanings across grade levels. Figure 2 below indicates that pre-school students mostly state *internal*-related force meanings. The elementary, middle and high school students mostly express *acquired*-related meanings. However, the *gravity and other* meanings scores are significantly higher for middle and high school students than the elementary school students' score.

Also, one way multivariate analysis of variance (MANOVA) was conducted to determine the effects of differences in grade levels (pre, elementary, middle, high) on the seven dependent variables, force meaning categories. Significant differences were found among the grade levels on force meaning categories, Wilks' lambda $\Lambda = .36$, $F(21, 195) = 3.90$, $p < .01$. The multivariate η^2 based on Wilks' lambda was strong, .29. Table 2 contains the means and standard deviations on the dependent variables for four grade levels.

Analyses of variances (ANOVA) on the dependent variables were conducted as follow-up tests to the MANOVA. In order to control Type I error, the Dunnett's *C* method was used. ANOVA on the *acquired, acquired/push-pull* and *gravity* scores were significant, $F(3, 74) = 4.02$, $p < .01$, $F(3, 74) = 12.82$, $p < .01$, $F(3, 74) = 9.16$, $p < .01$, respectively.

Table 1. DG&E sample coding schema for question set 1.

Force Meanings	Internal	Internal/ Move	Internal/ Acquired	Acquired	Acquired/ Push-Pull	Push-Pull	Gravity and Other
Set 1-Big vs. small stones standing on the ground.	-Force only on the big stone, but not due to air, gravity or ground. -Force on both stones but greater force on the big stone, but not due to air, gravity or ground.	-Force only on the big stone, but not due to air, gravity or ground. -Force on both stones but greater force on the big stone, but not due to air, gravity or ground.	-Force only on the big stone, but not due to air, gravity or ground. -Force on both stones but greater force on the big stone, but not due to air, gravity or ground.	-No force on any stone.	-No force on any stone. -Force only on the small stone but not due to gravity	-No force on any stone.	-Equal force on both stones -Force on both stones but greater force on the big stone.

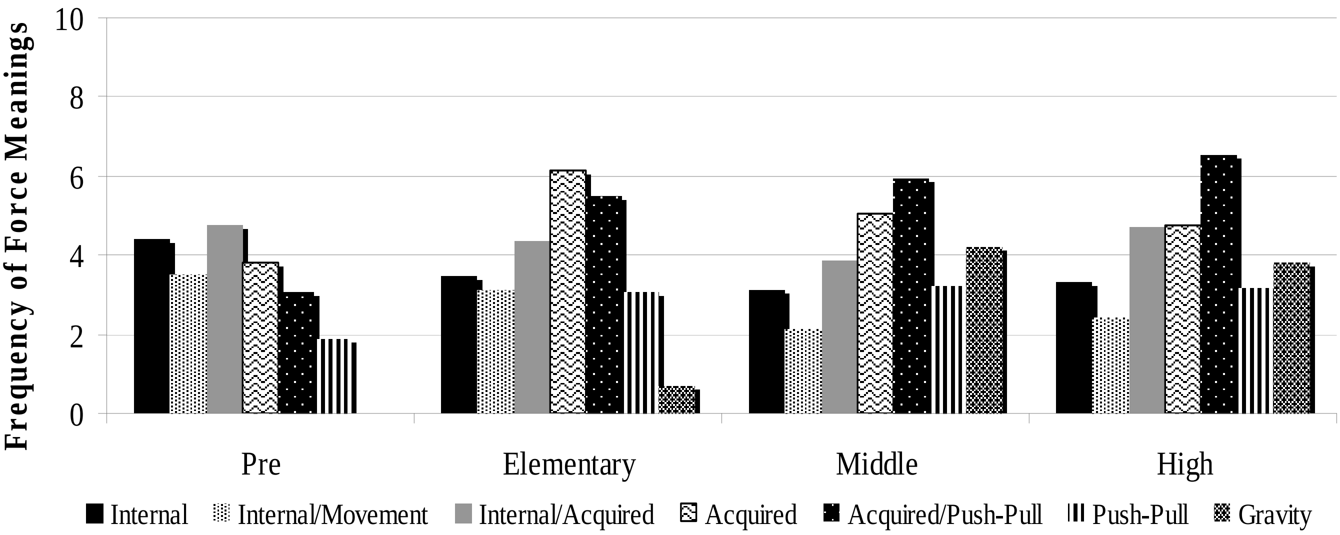


Figure 2. Turkish students' force meanings across grade level.

Table 2. Means and standard deviations for the force meaning categories

	Pre		Elementary		Middle		High	
Force Meanings	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Internal	4.40	3.07	3.47	1.94	3.13	2.90	3.32	2.15
Internal/Movement	3.53	1.81	3.12	2.36	2.13	1.41	2.42	1.54
Internal/Acquired	4.73	2.55	4.35	1.57	3.87	2.33	4.71	1.79
Acquired	3.80	1.32	6.12	2.49	5.07	1.94	4.74	1.78
Acquired/Push-Pull	3.07	1.87	5.47	2.06	5.93	1.71	6.55	1.67
Push-Pull	1.87	2.36	3.06	1.44	3.20	1.74	3.16	1.21
Gravity	.00	.00	.71	1.99	4.20	4.00	3.81	3.52

Post hoc analyses to the univariate ANOVA for the *acquired, acquired/push-pull* and *gravity* scores consisted of conducting pair-wise comparisons to find which grade level scores were higher on these categories. In terms of *acquired* category, the elementary school students got the higher force meaning scores than the pre and high school students. In terms of *acquired/push-pull* category, high, middle and elementary school students received significantly higher scores than the pre school students. Finally, in terms of *gravity* scores, high and middle school students received significantly higher scores than the elementary and pre school students. The middle and high school students were not significantly different from each other.

We also conducted a one way MANOVA to determine whether there is a difference between boys and girls for the force meaning categories. The MANOVA results was not significant at alpha level .05, Wilks' lambda Λ = .91, $F(7, 70) = .97$, *n.s.* Next, another one way MANOVA was conducted to evaluate the differences between students from Ankara and Gaziantep. Results indicate no significant differences between students from Ankara and Gaziantep in terms of force meaning categories, Wilks' lambda $\Lambda = .83$, $F(7, 70) = 2.02$, *n.s.* Therefore, we can conclude that either gender or the cities where students live were not the factors effecting for Turkish students force meanings.

Finally, we specifically analyzed the high school students' data from different content tracks. Twenty-three of the high school students in Gaziantep were from three different tracks as Math & Science, Math & Literature, and Social Sciences. A one way MANOVA was conducted to determine the effects of differences in tracks on the seven dependent variables, force meaning categories. Significant differences were found among the students from different tracks on force meaning categories, Wilks' lambda $\Lambda = .16$, $F(14, 28) = 2.97$, $p < .01$. The multivariate η^2 based on Wilks' lambda was quite strong, .59. The multivariate partial eta squared indicates 59% of multivariate variance of the dependent variables is associated with the track factor.

Separate ANOVAs also conducted on the dependent variables as follow-up tests to MANOVA. In order to control for Type I error across multiple ANOVAs, Holm's Sequential Bonferroni method was used. The ANOVA on *internal/acquired*, *push/pull* and *gravity* categories were significant.

Post-hoc analyses to the univariate ANOVAs for the scores from these three categories were conducted to find which track level affected the scores most strongly. First, math and science students received significantly higher scores than the math and literature students in the *internal/acquired* category. Second, social science students received significantly higher scores than the both math and science, and math and literature students for the *push/pull* ideas. Finally, in terms of *gravity* meaning scores, math and science students received significantly higher scores than the math and literature and the social science students.

Conclusion and Implications

The current study provides important results regarding Turkish students force meanings. The detailed analyses show that there are significant differences, in terms of force meanings, among grade levels but no difference by gender or the city students live. Also, the distributions of force meanings across grade levels indicate remarkable findings: First, force meanings are uniformly distributed for both middle and high school students. Especially, high school students demonstrated more consistent responses across question sets compared to the other group of students. Second, the pre-k students are spread across all of the force meanings categories except *gravity and other*. Third, the elementary school students are mostly clustered in the *acquired* related meanings. Finally, Turkish students are more evenly distributed across force meanings categories than U.S. students (diSessa, Gillespie, & Esterly, 2004; Özdemir & Clark 2009).

The school specific variations (e.g. Math & Science track) illustrate more important differences in terms of force meanings than the region specific variations (e.g., Ankara vs. Gaziantep). The results show that there are significant differences among students from different tracks in the high school. In order to understand the possible reasons, the analysis of the curricula (e.g., the degree in which students are exposed to the concepts of force in physics or science courses) and the school specific variations for different tracks (e.g., regular public schools versus anatolian high schools) are needed to be investigated and the study should be replicated with a much larger sample size.

No differences between pre-k students and middle-high school students in some categories are highly interesting results to look at. For example, observing no significant differences in terms of *internal* related force ideas between 5-6 years old pre-k students with no formal schooling and 13-16 years old middle and high school students with formal schooling (including many science courses) may illustrate thought provoking but disappointing implications for the implementation and the quality of Turkish science curriculum.

In addition to the limitations in the coding schema (read Özdemir & Clark 2009 for detailed analysis and limitations for the coding schema), the number of students involved in this study is not sufficient to generalize findings across Turkey. Also, all students in this study were attending urban schools; therefore, we do not have any data from rural or suburban schools. Finally, further analyses are needed to investigate the possible effects of culture and language in conceptualization of different science ideas.

Lastly, this study validates the general progressions outlined in the literature but suggests possibility of specific variations for Turkey. Overall, this study, with its relatively large sample size and in-depth interview questions, offers a significant contribution to the limited literature of Turkish students' force-meanings.

References

Anderson, C. W. (2007). Perspectives on Science Learning. In S. K. Abell & N. G. Lederman (Eds). *Handbook of research on science education*. (pp. 3-31). Mahwah, NJ: Lawrence Erlbaum.

Chi, M.T.H. (2005). Commonsense conceptions of emergent processes: Why some misconceptions are robust. *Journal of the Learning Sciences, 14*(2), 161-199.

Clark, D. B., D'Angelo, C., & Schleigh, S. & Menekse, M. (2008). International comparison of knowledge structure coherence: Cultural, semantic, and educational differences. In *Proceedings of the Eighth International Conference of the Learning Sciences*. Utrecht, Netherlands.

Clark, D. B., Menekse, M., Özdemir, G., & D'Angelo, C. (in preparation). Comparison of knowledge structure coherence and force meanings across sites in Turkey.

Clement, J. (1982) Students' preconceptions in introductory mechanics. *American Journal of Physics*, *50*, 66-71.

Cooke, N. J., & Breedin, S. D. (1994). Constructing na¨ıve theories of motion on the fly. *Memory and Cognition*, *22*, 474–493.

diSessa, A. A., Gillespie, N., & Esterly, J. (2004). Coherence versus fragmentation in the development of the concept of force. *Cognitive Science*, 28, 843-900.

Duit, R. (2007). *Bibliography STCSE –Students' and Teachers' Conceptions and Science Education*. Kiel, Germany: IPN Leibniz Institute for Science Education.

Gunstone, R., & Watts, M. (1985). Force and motion. In R. Driver, E. Guesne, & A. Tiberghien (Eds.), *Children's ideas in science* (pp. 85-104). UK: Open University Press.

Halloun, I., & Hestenes, D. (1985). The initial state of college physics students. *American Journal of Physics, 53*, 1043-1055.

Ioannides, C., & Vosniadou, S. (2002). The changing meanings of force. *Cognitive Science Quarterly*, 2(1), 5-62.

McDermott, L. (1984). Research on conceptual understanding of physics. *Physics Today*, *37*, 24-32.

Özdemir, G. & Clark, D. B. (2009). Coherence of Turkish students' understanding of force. *Journal of Research on Science Teaching, 46*(5), 570-596

Watts, D. M., & Zylbersztajn, A. (1981). A survey of some children's ideas about force. *Physics Education, 15*, 360-365.

WHAT CHILDREN AGED FIVE TO SEVEN YEARS KNOW ABOUT SEXUAL REPRODUCTION

Cristina Martínez-Losada
A Coruña University. Spain

Susana García-Barros
A Coruña University. Spain

María Garrido
A Coruña University. Spain

Abstract

By means of individual interviews, this study looks into what 315 children at the initial educational levels in Spain know about certain reproduction-related aspects. Moreover, a small group was monitored for a three-year period with the aim of learning how these particular children's knowledge evolved as they became older. The children recognise the way in which closely specimens reproduce. In general, they perceive the intervention of the male more easily where viviparous example is concerned and they see the transformation of the individual before birth more clearly where oviparous specimens are involved, though, the knowledge about this matter is greater in the older pupils. The personal ideas of these children evolve over time, especially as regards viviparous example.

Introduction

Reproduction is an important life function for the characterization of life. Therefore, beginning to study even it from the most basic of educational levels allows children to elaborate an increasingly more adequate living being concept (García Rovira, 2005). On the other hand, conceptualising it adequately is an essential requirement in order to come to understand more complex concepts such as heredity, adaptation, evolution…, which are studied in Secondary Education and are particularly difficult for pupils. (Wood-Robinson, 1994; Andersson & Wallin, 2006).

All through Primary Education, pupils are expected to develop increasingly more complex and systemic ideas about the reproduction function. More specifically, the Spanish Curriculum includes aspects such as the types of human reproduction and that of other living creatures too, their characteristics, etc. However, understanding these life functions is not at all easy for pupils who have had preconceived ideas, basically taken from their family context, ever since they were little children (Ramos, 2008). The teacher must be aware that such ideas exist as well as of the need to help pupils to further them.

Based on what has pointed out thus far, this paper intends to find out what children at the initial educational stages know about oviparous/viviparous reproduction and how their ideas about it develop as they become older.

Rationale

There are different interpretations about how the youngest children acquire their ideas about biology, and some authors like Carey (1985) give them a mainly psychological origin, whereas others (Mintzes, 1984; Hatano & Inagaki, 1997) consider children to have intuitive knowledge about the way the human body works which they can extrapolate to other animals.

Several studies aimed at establishing what ideas very young children have about the digestive system (Carvalho *et al.*, 2004; Cakici, 2005), the respiratory system (Cuthbert, 2000) or about the organs and systems of the human body and that of other animals, in general, have been undertaken. (Reiss & Tunnicliffe, 2001a; 2001b). However, studies aimed at establishing what ideas children have about reproduction are scarcer. In this case, preformist type explanations about the origin of the new individual – similar to those used by Science in the past – have been identified (Bernstein & Cowan, 1975; Carey, 1985; Goldman & Goldman, 1982). Furthermore, children may have a different reproductive model for human beings than for oviparous animals (Giordan & De Vecchi, 1988; Russell & Waat, 1990).

On the other hand, the fact that the teaching work during the first few years focuses on the aspects differentiating both types of reproduction does not make it easier for pupils to acquire a unitary idea of what reproduction really is (Martínez-Losada *et a*l, 2008).

Methodology

Three hundred and fifteen children have been interviewed, individually (74 of them five years old, 101 of them six years old and 91 of them seven years old). In addition to this, 32 children (from five to seven years of age) were monitored for three years.

The interview consists in the following questions, associated to specific illustrations:

- Could you identify how the human being, dog, lion, duck, hen, vulture, fish, frog, cocodrile, turtle and bee?

The purpose of this conversation is to learn whether the children are able to identify the sort of reproduction involved (oviparous/viviparous) where certain types of examples are concerned

- Does the father take part in human reproduction?. Does the father take part in ducks reproduction?. Dou you know how to draw the new individual before birth?.

The purpose of this conversation is to learn whether the children are able to recognise the intervention of the male element in human reproduction (the prototype of viviparous animals) and ducks reproduction (the prototype of oviparous animals) and the children are able to depict the individual's transformation before birth.

To make the result analysis process easier, the drawings made by the children were divide into different groups depending on: a) the transformation of the individual is not recognised; b) the transformation of the individual is recognised.

In Figure 1, representative drawings for each group are shown.

Findings

All of the children recognise human beings and mammals as viviparous animals and they also recognise the birds found in their everyday environment as oviparous ones (figure 2). These children recognise the oviparous reproduction of other animals to a lesser extent, with the exception of the bee, although, they improve their ability to recognise it, progressively, as they become older.

Taking a closer look at the children's knowledge about the reproduction of specific specimens (the human being and the duck), the children's answers have been grouped into five different categories: (A) neither the transformation of the individual nor the intervention of the male element is recognised; (B_1) the intervention of the male element alone is recognised; (B_2) the transformation of the individual alone is recognised; (C) both aspects are recognised and (DNK) does not know.

Different sorts of categories have been detected depending on the child's age and on the example (figure 3). In the case of the duck, about 50% of the answers given by the subjects of different ages were either type A or type B_2 (recognising the transformation of the individual). The first answers were the most frequent ones in the group of five-year-olds, and the second ones in the other groups. The type C answers (recognising the transformation of the individual as well as the intervention of the male) are not found at all in the group of five-year-olds and are hardly found in the other two.

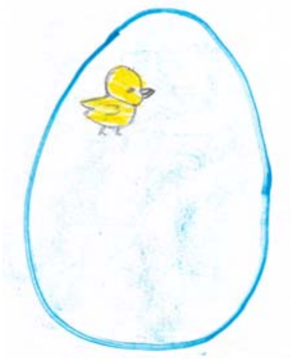
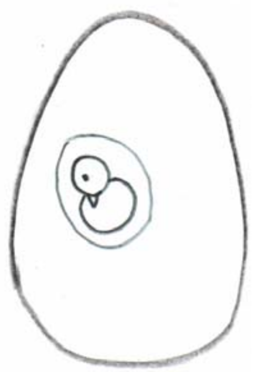
Group	Duck	Human
The transformation of the individual is not recognized		
The transformation of the individual is recognized		

Figure 1. Examples of drawings made by children.

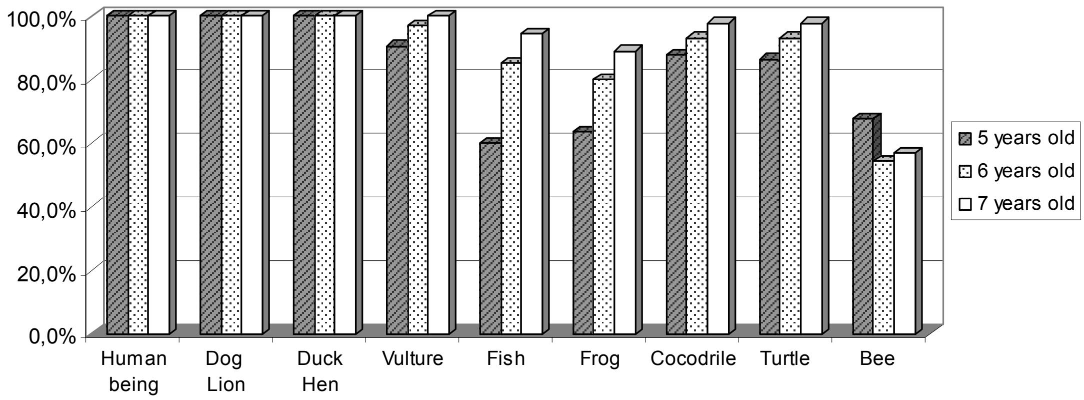


Figure 2. Percentages of little boys/girls who recognise the way different species reproduce.

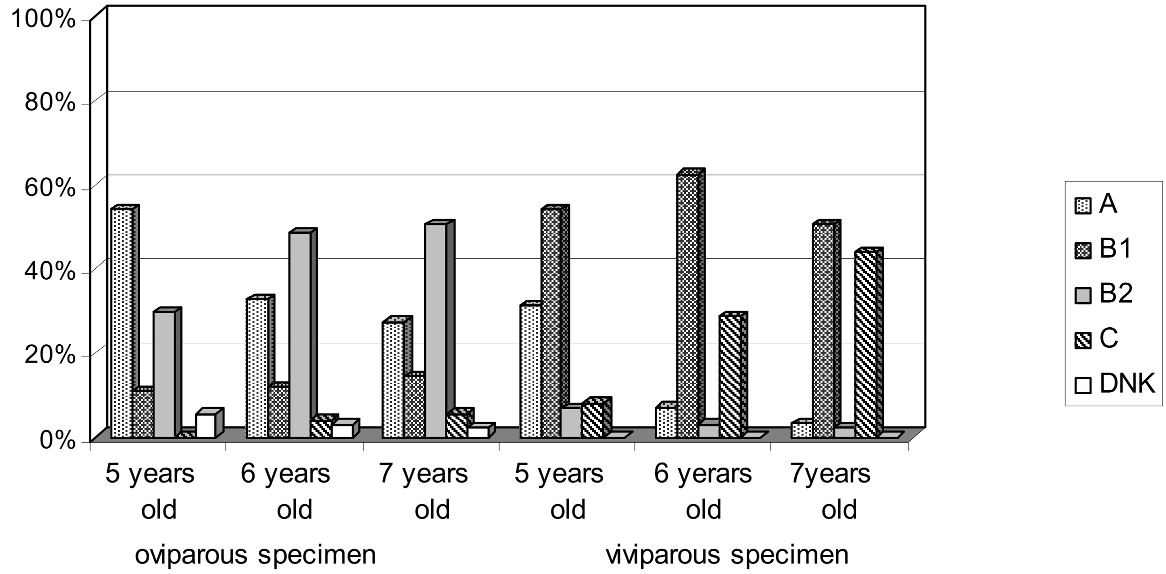


Figure 3. What children know about the reproduction of oviparous and viviparous specimens.

In table 1 and table 2, the development of the children's ideas about this matter over three consecutive years is shown.

With regard to the reproduction of the duck (table 1), 23 out of 32 children evolve, although their final answers were never type C. Most frequently, the children evolve from A → B_2 answers (12 children) and A → B_1 ones (7 children), which means one change only with regard to their previous answer and they usually occur in the five-six year interval (17 children).

As regards the reproduction of human beings, 24 out of 32 children evolve, and the most frequently found answer is the type C one (20 children). Six children gradually modify their answers, continuously, as time goes by, and they all evolve from the most primitive type of answer (A) to the most elaborate one (C). In the rest of the cases, the evolution takes place in only one of the age intervals, namely the one comprehending children aged five to six (13 children). This may mean one single change with respect to the children's previous answer – B_1 → C (six children), A → B_1 (4 children) and B_2 → C (2 children) – or a greater improvement – A → C (6 children)-.

Table 1. Ideas which children aged five to seven have about the oviparous specimen.

Ideas of children (5 → 6 → 7 years)		Identifying the children	Total
They evolve (n=23)	DNK → A → A	32	1
	DNK → B_1 → B_1	34	1
	DNK → A → B_2	6	1
	A → A → B_1	19, 25	2
	A → B_1 → B_1	2, 12,16, 28, 29	5
	A → A → B_2	3,15,27	3
	A → B_2 → B_2	4, 5, 9, 10, 11, 14, 17, 31, 36	9
	B_1 → B_2 → B_2	21	1
They do not evolve (n=9)	A → A → A	22, 26	2
	B_1 → B_1 → B_1	18	1
	B_2 → A → A	20	1
	B_2 → A → B_2	13, 30	2
	B_2 → B_1 → B_2	8	1
	B_2 → B_2 → B_2	1, 23	2

Conclusions and Implications

The children are able to identify the way in which closely animals reproduce. Then, they extend it to other animals, and this is consistent with the emphasis placed on this particular matter by the Spanish Educational System at these initial levels.

Generally speaking, it is easier for the children to perceive the intervention of the male element in the viviparous example, and the transformation of the individual in the case of the oviparous specimen. Furthermore, the knowledge about reproduction is greater in the older pupils, above all where the viviparous animal is concerned. The individualised analysis of what children belonging to a small group know about sexual reproduction confirms what has previously been stated. Thus, most of them evolve during this period, though greater changes are seen in connection with the viviparous example than with the oviparous one.

Education ought to place bigger emphasis on the oviparous reproduction and, more specifically, on the intervention of the male element, because this phenomenon creates a greater degree of difficulty. The importance of being able to recognise the intervention of the male in reproduction has been highlighted in other research works, because it accounts for the differences which are seen between children and parents (Pujol, 2003; García Barros & Martínez Losada, 2006).What has been shown so far will, no doubt, help to favour a more unitary idea about sexual reproduction from the initial educational levels, thus enabling children to go beyond the nearly exclusive emphasis which has been given to establishing differences.

Table 2. Ideas which children aged five to seven have about the viviparous specimen.

Ideas of children 5 → 6 → 7 years old		Identifying the children	Nº of children
They evolve (n=23)	A → A → B_1	23	1
	A → B_1 → B_1	1, 22, 29	3
	A → A → C	4	1
	A → B_1 → C	8, 11, 14, 20, 30	5
	A → B_2 → C	27	1
	A → C → C	16, 17, 19, 28, 31	5
	B_1 → B_1 → C	2, 3, 21	3
	B_1 → C → C	5, 25, 36	3
	B_2→ C → C	6, 18	2
They do not evolve (n=9)	B_1→ B_1 → B_1	9, 10, 12, 13, 15, 32, 34	7
	B_2→ B_2 → B_2	26	1

References

Andersson B. & Wallin A. (2006). On Developing Content-orinted Teories Taking Biological Evolution as an Example. *International Journal of Science Education* **28,** 673-695.
Bernstein A. C., & Cowan P. A. (1975). Children's concepts of how people get babies. *Child Development* 46, 77-91
Cakici Y. (2005). Exploring Turkish upper primary level pupils' understanding of digestion. *International Journal of Science Education* **27,** 79-100.
Carey S. (1985). *Conceptual change in childhood.* The MIT Press, Cambrige MA.
Carvalho G. S, Silva R., Lima N. & Coquet E. (2004). Portuguese primary School children´s conceptions about digestion: identification of learning obstacles. *International Journal of Science Education* **26,** 1111-1130.
Cuthbert A. J. (2000). Do children have a holistic view of their internal body maps? *School Science Review* **82 (299),** 25-32.
García Barros S. & Martínez Losada C. (2006). La reproducción. responsable de promover diferencias en la semejanza. *Padres y Maestros* **302,** 23-27.
García Rovira M. P. (2005). Los modelos como organizadores del currículo en Biología. *Enseñanza de las Ciencias* Nº extra. VII congreso , 1-5.
Giordan A. & De Vecchi C. (1988). *Los orígenes del Saber*. DIADA Editores, Sevilla.
Goldman, R. J., & Goldman, J. D. G. (1982). How children perceive the origin of babies and the roles of mothers and fathers in procreation: a cross-national study. *Child Development* 53, 491-504.
Hatano G. & Inagaki K. (1997). Qualitative changes in intuitive biology. *European Journal of Psichology of Education* **21,** 11-130.
Martínez-Losada, C. García-Barros, S. & Garrido, M. (2008). Los seres vivos en las actividades de enseñanza que se emplean en los niveles educativos iniciales. *Actas de los XXIII Encuentros de Didáctica de las Ciencias Experimentales*, pp. 1058-1068. Almería.

Mintzes J. J. (1984). Naive theories in biology: children's concepts of the human body. *School Science and Mathematics* **84,** 548-555.

Prokop, P., Prokop, M., Tunnicliffe, S. D. & Diran, C. (2007). Children´s ideas of animal´s internal structures. *Journal of Biological Education*, **41(2**), 62-67

Pujol R. M. (2003). *Didáctica de las ciencias en la Educación Primaria*. Síntesis educación, Madrid.

Ramos, J. (2008). Algunas teorías sobre distintos aspectos del cuerpo humano. Implicaciones didácticas. Alambique, 58, pp. 77-88.

Reiss M. J. & Tunnicliffe S. D. (2001a). Student´s understandings of human organs and organ systems. *Research in Science Education* **31,** 383-399.

Reiss M. J. & Tunnicliffe S. D. (2001b). Students´understandings about organs and organ systems in differerent animals. In *III Conference of European Researchers in Didactic of Biology*, ed. García Rodeja I., Harms U., JiménezAleixandre M. P., pp. 113-125. Universidad de Santiago de Compostela, Santiago de Compostela.

Russell T. & Waat D. (1990). SPACE Research Report. *Journal of Biological Education* **35,** 61-65.

Wood-Robinson C. (1994). Joung peoples's inheritance and evolution. *Studies in Science Education* **24,** 29-47.

A PHENOMENOGRAPHIC FRAMEWORK FOR ANALYSING UNIVERSITY STUDENTS' UNDERSTANDING OF ELECTROMAGNETIC INDUCTION

Kristina Zuza
Univeristy of the Basque Country

Jenaro Guisasola
Univeristy of the Basque Country

José Manuel Almudí
Univeristy of the Basque Country

Abstract

From a socio-constructivist point of view of teaching knowledge of students' forms of reasoning is essential in the task of reconstructing the aims and contents of teaching. Teaching-learning of the explanatory theory of Electromagnetic Induction (EI) phenomena is a little researched pedagogical problem, if compared with other physics subjects. Teaching EI in a physics programme for first-year science and engineering students is interesting for many reasons. For this investigation 85 students were chosen on an initial Industrial Electronic Engineering degree course at the University of the Basque Country and we used qualitative questions which focussed more on explanations than on obtaining a result by applying a formula. In the five questions students were asked to explain EI phenomena brought about by variable magnetic fields or by the movement of a circuit or part thereof in a stationary magnetic field, the great majority of students used an explanatory model based on the variation in magnetic flow and only a minority used a model based on the forces acting on the charges even when they had been clearly asked to do so. This research provides insights into students understanding of EI which might be use for designing teaching sequences.

Background and purpose

From a socio-constructivist point of view of teaching (Leach & Scott 2003), knowledge of students' forms of reasoning is essential in the task of reconstructing the aims and contents of teaching. In the field of Electricity and Magnetism, a lot of research work has shown that students get poor teaching in basic matters of electrostatics, continuous current circuits or magnetic fields, among other things. However, teaching-learning of the explanatory theory of Electromagnetic Induction (EI) phenomena is a little researched pedagogical problem, if compared with other physics subjects. In this work we will attempt to answer the following research question: What conceptions and forms of reasoning do first-course science and engineering students use to explain electromagnetic induction phenomena?

Teaching electromagnetic induction (EI) in a physics programme for first-year science and engineering students is interesting for the following reasons: a) It is a subject in which different laws and concepts in the electrical and magnetic field are dealt with jointly; b) The existence of previous studies which show students' difficulties in analysing concepts such as magnetic flow or Faraday's law; c) A correct interpretation of EI phenomena allows people to take informed decisions on many applications of EI in everyday life (induction cookers, electric motors, etc.).

Rationale

The investigation of students' ideas and forms of reasoning has been made by using a phenomenographic methodology (Buck et al., 2003, Marton and Booth, 1997). The aim of phenomenograhic research is to obtain a group of categories which describe the qualitative variations in the forms in which participants (e.g. students) experience, interpret, understand, perceive or conceptualise a subject for study, an activity or a phenomenon (e.g. electromagnetic induction).

Previous phenomenograhic research used the individual interview as an initial source of data. In the work presented here, the main source of data has been the written replies to a questionnaire. In principle, Marton and Booth (1997) do not show that there are impediments to using questionnaires as a source of data or other types of techniques which in some way act as an expression of how people perceive and experience events. In a phenomenographic study of students' conceptions of gravity, Sharman et al (2004), use the written replies to a question as a source of data. They point out (p. 270) "In phenomenography, particularly when we are looking at written responses rather than interviewing, we prefer a large sample size in order to pick out the smaller categories". In our study we consider that a sample of 85 answers from two classes of engineering students is sufficient to meet the conditions imposed in research using this kind of approach.

Methods

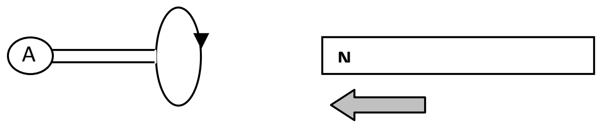
The 85 students were chosen at random from two groups on an initial Industrial Electronic Engineering degree course at the University of the Basque Country. These students were taught on the Physics Course by two professors with a minimum of eight years' teaching experience in the Department of Applied Physics. The students did the questionnaire one month after having been taught EI, when they sat the final examination on the subject.

As recommended by White and Gunstone (1992), we used qualitative questions which focussed more on explanations than on obtaining a result by applying a formula. The answers were examined independently by three members of the research group who looked for similarities and differences between the explanations, selecting significant statements and comparing these statements in order to obtain cases of agreement or variations and afterwards grouping them in categories. After the initial categories had been established the researchers met to discuss these categories and review them until a consensus on final categories was reached. With these categories, the questionnaires were checked again to determine whether the categories were descriptive and indicative enough of the data. This analysis process is consistent with phenomenograhic analysis, as Marton (1981, p. 43) claims "definition for categories are tested against the data, adjusted, retested, and adjusted again".

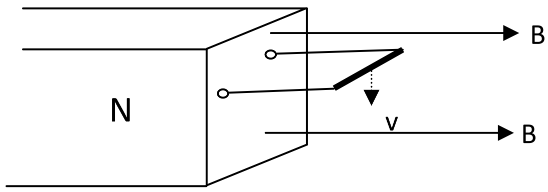
Results

In the five questions students were asked to explain electromagnetic induction phenomena brought about by variable magnetic fields (Q1 and Q2) or by the movement of a circuit or part thereof in a stationary magnetic field (Q3, Q4 and Q5). The questions will be explained in detail in the communication; here we give questions Q2 and Q4 as examples.

Q2. - We have a magnet which moves towards a conducting loop which is at rest in respect of our observation (see figure); the ammeter can record the passage of an intensity through the loop at any time during the approach process. As you studied the electric current in the conducting spiral is due to an electrical force associated with an electrical field. Explain how this electrical field appears in the spiral and the nature thereof.



Q4. - A U-shaped wire is sliding along a magnet as shown in the figure maintaining the angle in respect of the magnetic field. Bearing in mind that both the wire and the magnet are conductors, is there an induction phenomenon in the wire? Justify your answers.



In the case of EI phenomena similar to those analysed in class, the great majority of students used an explanatory model based on the variation in magnetic flow (between 50% and 75%) in electromagnetic phenomena produced by a variable magnetic field (Q1 and Q2), and in those caused by the movement of the circuit or part thereof (Q3, Q4 y Q5). Only a minority used a model based on the forces acting on the charges even when they had been clearly asked to do so (3.5% in Q2 y 4% en Q3). The results obtained, as well as the types of answers, will be explained in the communication.

Conclusions and Implications

After the different types of answers had been identified; the next task was to interpret them to define the phenomenograhic categories. We defined the categories on the basis of the data obtained and the criteria of Marton y Booth (1997, p. 307). Four explanatory categories have been identified. The first category, "descriptive-memoristic", (15%-20% of replies) includes explanations which only give descriptions of the phenomenon or state meaningless electromagnetic concepts and laws to explain the phenomenon in question.

We have called the second category 'Inherent to the nature of the magnetic field or the electrical current' (20% of answers). This category includes the answers which attribute electromagnetic induction to the presence of a magnetic field, whether variable in time or not, or to the presence of an electric current. The third category, which we have called "variation in magnetic flow" (50%-75%), is characterised by analysing any EI phenomena by using Faraday's law and the variation in magnetic flow from a macroscopic perspective, However, many students in this category made an incorrect analysis when trying to analyse EI phenomena caused by the movement of the circuit or a conductor (question Q5) by confusing the surface of the conductor with the surface swept by the movement of the conductor (more than 40%). The last category, which we have called "Faraday-Lorentz" (5% of answers), includes answers which explain EI phenomena both in terms of field (Faraday's Law) and in terms of actions of the field (Lorentz's Law).

The important thing about our result is not that it shows poor learning of the EI theory (this is a well established fact) but that it shows some components of students' failure to understand; for example, students did not differentiate between macroscopic models based on active fields and microscopic models based on the forces exerted by these fields. The results obtained do not claim that the categories detected can be used to describe all students' conceptions about EI, or that each student falls into one of the explanatory categories when answering a question on electromagnetic induction. However, these categories bear similarities with the results obtained in other research on students' ideas about EI phenomena (Loftus 1996), although a longitudinal study with students from other courses will be necessary to see whether these descriptive categories are maintained throughout teaching.

This research provides insights into students understanding of EI which might be use for designing teaching sequences. This question will be the subject of our future researches.

Students' difficulties seem strongly linked to the absence of an analysis of the work carried out on the circuit and its energetic balance. In this regard, most 3rd course physics students still do not clearly understand the usefulness of concepts of potential difference and fem. The difficulties found and the contents of the analysis of the physics show the necessity to provide students with opportunities for them to reflect on the role played by the concepts of potential difference and electromotive force in the energetic model which explains the movement of the current in a simple electrical circuit. It will be necessary to design tasks and problems which lead to it being understood that the difference between the concepts of electromotive force and potential difference is given to measure different kinds of actions produced by radically different causes; the former due to non-conservative causes and the latter to conservative forces. This involves knowing that electromotive force is a quantity which quantifies a transfer of energy (from a battery to the charges of a circuit) associated with a non-conservative action.

References

Buck P., et al (2003). On the methodology of 'phenomenography' as a science education research tool, En D. Psillos et al., (Edts.) *Science Education Research in the Knowledge- based Society* , p.31-41. Kluwer Academic Publishers.
Leach J. & Scott P. (2003). Individual and sociocultural views of learning in science education. *Science and Education*, 12(1), 91-113.
Loftus, M. 1996. Student's' ideas about electromagnetism. SSR 77, 280-291.
Marton, F. (1981). Phenomenography-Describing conceptions of the world around us, *Instructional Science*, 10, pp. 177-200
Marton, F. y Booth, S. (1997). *Learning and awareness*, Mahwah, NJ: Lawrence Erlbaum Associates Publishers
Sharman, M., Millar, R., Smith, A. and Sefton, I. (2004) Students' understanding of gravity in an orbiting space-ship, *Research in Science Education* 34, 267-289.
White, R. and Gunstone R., *Probing Understanding*. The Palmer Press. London.

STUDENTS' MOTIVATION ON LEARNING MATERIAL SCIENCE TEACHING MODULES IN FIVE COUNTRIES

J. Lavonen, R. Byman, A. Loukomies, V. Meisalo
University of Helsinki, Finland

C. Constantinou, T. Kyratsi, N. Papadouris
University of Cyprus, Cyprus

D. Couso, M.I. Hernandez, R. Pinto
University Autonoma de Barcelona, Spain

E. Hatzikraniotis, M. Kallery, E. Petridou, D. Psillos
Aristotle University of Thessaloniki, Greece

P. Kariotoglou, D. Pnevmatikos, A. Spyrtou
University of Western Macedonia at Florina, Greece

S. Lombardi, G. Monroy, I. Testa
University "Federico II" of Naples, Italy

Abstract

Teaching modules for comprehensive schools in specific domains related to materials science were designed and developed in five countries participating in the EU's 6th framework project MaterialsScience (MS). The motivational properties of the designed modules were analysed through content analysis in the framework of Self-Determination Theory (SDT). Teaching experiments, including science inquiry activities, were organised in each participating country with at least 30 students taking part. Motivational features of the science activities in general and MS module activities were evaluated by the students with an Evaluation of Science Inquiry Activities – Questionnaire (ESIAQ). The ESIAQ is a multidimensional measurement device, developed in the framework of SDT. The motivational features of the MS module activities and science activities, in general, were evaluated very similarly in each country by students with ESIAQ. The most motivating feature of the MS module's activities in all countries was the support for student interest and the least motivating feature was the support for the feeling of student autonomy.

Introduction

Students' motivation towards and interest in science and technology have been intensively researched since the 1960s. It is known that science in general is quite interesting for students, but most students, especially girls, do not find school science and technology or careers and occupations in those fields interesting (Osborne, Simon & Collins, 2003)

The objective of the EU 6th framework project, *MaterialsScience*[1] (MS), has been to design in five countries MS teaching modules for comprehensive schools in specific domains related to materials science. These modules aimed to emphasise the inquiry based science learning, the active student engagement, the motivation and the collaborative learning. In this paper, the motivational aspects of the developed activities are analysed.

The concept 'motivation' in the MS-project has been used to describe the factors within an individual (including an interaction with the environment) which arouse, maintain and channel behaviour towards the aims of the MS teaching modules. Conceptualisation of motivation is based on the *Self-Determination Theory* (SDT). Central to SDT is the concept of basic psychological needs that are assumed to be innate and universal. These needs are the *need for autonomy*, the *need for competence*, and the *need for relatedness (need to belong to a group)*. Furthermore, the *interest* of the student in a learning activity have an effect to motivation (Deci & Ryan, 2004). Self-determined learning occurs when a learning activity itself supports fulfilment of the basic psychological needs or development of interest. Both, the features of a learning activity and behaviour of a teacher, could potentially increase the motivation of a learner.

In the MS-project six different modules, with different motivational features, have been developed for upper primary and lower secondary level. Teaching experiments, including science inquiry activities have been organised in each participating country. In Cyprus students were engaged in small groups in a series of inquiry-based activities intended to help them develop an explanatory framework for electromagnetic phenomena and, secondly, in the design and construction of a train that implemented levitation and propulsion mechanisms. In Finland, the students learnt about the manufacturing and use of materials as well as material science careers within an industry site visit. The students planned and learned during the visit in small groups. In Greece (Florina) the students (10-11-year olds) learnt about the density as an "identity" of different materials, as well as about how to predict floating/sinking phenomena. The distinction and control of variables as well as the reflection upon this procedure and models and modeling were used as elements of scientific inquiry. In Greece (Thessaloniki) students learned about thermal properties of materials through inquiry activities. In Italy the students have been engaged in experiments and computer-based simulations about guide lights and optical fibres. During the proposed activities, the students worked in small groups. In Spain, the selected inquiry activity chosen for assessing its motivating features refers to the analysis of the acoustic behaviour of different materials by carrying out experiments with data capture systems to measure the sound attenuation. Its main aim was to promote the development of coherent conceptual understanding on acoustic properties of materials and an awareness of the nature of science by engaging students in active intellectual tasks and experimental work in collaborative settings. It is obvious, that each module could fulfil different basic needs of the learners.

The research questions were:

1. What motivational features are there in the designed six material science modules?

2. What features a) in school science activities, in general, b) in science activities of the MS modules motivate students towards science activities in each participating country?

Methods

Content analysis is used to analyze the motivational features of each MS module. One local researcher in each country analysed the motivational features of their MS module and categorised the motivational features according to four categories: *autonomy-supporting activities, support for students' feeling of competency, support for students' social relatedness, and support for student interest.* The categories that was used in the analysis are based on SDT and, therefore, the analysis is theory driven.

[1] University-school partnerships for the design and implementation of research-based ICT-enhanced modules on Material Properties, SAS6-CT-2006-042942-Material Science (042942)

At least 30 students participated to each teaching experiment based on the MS modules. *An adapted version of the Evaluation of Science Inquiry Activities Questionnaire* (*ESIAQ*) was used to collect data in order to investigate our second research question. The *ESIAQ* is based on Intrinsic Motivation Inventory (IMI)[1] which is a multidimensional measurement device intended to assess participants' subjective experience towards to a target activity (here a science inquiry activity). The IMI has been used in several experiments and settings related to intrinsic motivation and self-regulation (e.g., Deci, Eghrari, Patrick, & Leone, 1994), and has a strong support for its validity (*see* McAuley, Duncan, and Tammen, 1987). The *ESIAQ* uses a seven point likert scale (1 = item in my case not at all true … 7 = item in my case very true). Five (*perceived autonomy/choice, perceived competence, support for relatedness, enjoyment and value or usefulness*,) of the seven subscales of the *ESIAQ* was used in the present study (subscales *effort,* and *felt pressure or tension w*ere excluded in this study). The general criteria for the inclusion of items on subscales was a low factor loading (< .600) on the appropriate subscale, and no cross loadings above .400. A number of five to seven items were included in each subscale.

Different cultural or educational contexts (i.e. teachers' everyday practices) usually have different impact to students' motivation and interest towards different targets (OECD, 2007). Thus, within, rather than between modules, differences could be more informative about the features that motivate students towards science. In each country, participants were tested with *ESIAQ* twice: before (measurement of the motivational features in general) and immediately after teaching experiments (measurement of the motivational features of the MS modules). The data acquired by the *ESIAQ* before and after the teaching experiment will be compared in each country – not between countries. This was done because there are differences in cultural context between countries. Moreover, the teachers and their way to teach vary within one country and between countries. Based on the comparisons in a country, it could be possible to recognise the motivational features of designed MS modules.

Results

One researcher in each participating country did a content analysis of the developed MS teaching modules based on the five categories conducted from *SDT*. These categories and the six most important motivational features of each country's MS modules are presented below (CY = Cyprus, FI= Finland, GRF = Greece (Florina), GRT = Greece (Thessaloniki), IT = Italy, ES = Spain):

1 autonomy-supporting activities or support for choices
- the student-centred learning methods or use of ICT (FI, ES, GRF, GRT, CY);
- co-planning of the learning activities or students have choices of how to study (FI, IT)

2 support for students' feeling of competency
- choice of tasks, which are possible for the student to solve (ES,IT,CY, GRF);
- choice and use of constructive evaluation methods (IT, CY);

3 support for students' social relatedness
- choice of activities and ICT use which help students to feel close to peers (FI, ES, GRF. GRT);
- the feeling that the students can trust each other (IT,);

4 support for interest
- to awaken feeling related components of interest (enjoyment) (FI, ES, IT, GRF, GRT, CY);
- to awaken value related components of interest (usefulness) (FI, ES, GRF, GRT, CY);
- material science content and context (FI, ES, IT, GRF, GRT, CY).

Comparisons of students' evaluations of motivational features of the science activities, in general (1), and of the some crucial activities of the MS module (2), researched with *ESIAQ* in each country, are presented in Table 1. Five mean scores, one for each of the five subscales were created for the participants' motivation for the common (M_1 and $S.D._1$), and five mean scores for the MS modules activities (M_2 and $S.D._2$). For the five subscales (the sum variables of subscale items) means and Standard deviations (and) in

[1] http://www.psych.rochester.edu/SDT/measures/intrins.html

each participating country was calculated using SPSS. Comparison of motivational features of science activities in general and MS module activities ($M_2 - M_1$) was calculated through the paired-samples *t*-test procedure using SPSS.

Conclusions and Implications

In all countries, the motivational features of both the MS module activities and science activities, in general, were evaluated in most of the subscales similarly by students. The students' evaluations lie typically over the middle of the scale with the most motivating feature of the MS module activities, in most of the countries, especially in Greece (Florina) and in Italy, the value or usefulness for the students' learning and for their future life.

However, students experienced higher motivation on the MS module activities only in 'perceived autonomy' (Greece, Florina) and in 'interest/ value or usefulness of the activities' subscales (Italy). Contrary, students evaluated MS module activities supporting significantly less their 'feeling of autonomy' (Italy), 'interest / value or usefulness' (Greece, Thessaloniki and Spain) and 'perceived competence' (Spain) than common science activities in general.

Three were the most important findings from this project concerning the students' motivation. First, it is difficult to positively change all the motivational features due the short time intervention. It is quite possible that students will show signs of corrosion of their motivation towards their target. This destabilization and regression is not always a 'bad sign' in the means of the development. Alternatively, they could be seen as a sign of the reorganization of the aspects in matter. In this case it is important to compare how students with different motivational profiles change their attitude towards science activities. Second, the different motivational features and the priorities each country had when they organised their MS module, probably, had different results on the student's motivation. For example, the positive significant differences on perceived autonomy in the Greek (Florina) experiment in combination with the Greek traditional teaching that do not emphasize students' autonomy had an effective result on this aspect. Third, teachers should not be disappointed if they experience difficulties to change their students' motivation towards science. Changing students' motivation should be a long term procedure with improvements and, probably, regressions.

There are several limitations of this study. For example, we are taken for granted that all the teachers implemented the module using the teaching strategies described in the module descriptions. However, we know based on our observations that there were several differences in their interpretations of what was expected from them, and therefore, in the way they implemented. For example, in the Spanish context, data on motivation was collected from 4 different groups of students and therefore, 4 different teachers. These teachers supported students' autonomy very differently. Moreover, due to their different educational context it is quite possible that students experienced the activities and questionnaire questions in a different way and a deeper assessment (through the interviews) is needed to clarify their motivational experiences. Finally, we could accept that some of the intentions on motivation of the designed modules have not been fulfilled and therefore, some refinements of certain motivational aspects need to be done and more emphasis on these aspects should be explicit for teachers who implement the MS modules in their science classes.

Table 1. Means and Standard deviations for motivation subscales based on students' evaluations in each participating country and comparison of motivational features of science activities in general and MS module activities.

Country	Motivational features of the science activities in general and MS module activities	*N*	Science activities M_1	Science activities $S.D._1$	MS module M_2	MS module $S.D._2$	$M_2 - M_1$	t *)
Cyprus	Perceived autonomy/choice[1]	124	3.69	1.22	3.61	1.30	-0.08	0.7^{ns}
	Perceived competence[2]	124	4.20	1.13	4.18	1.12	-0.03	0.2^{ns}
	Support for relatedness[3]	124	4.20	1.17	4.07	1.16	-0.13	1.2^{ns}
	Interest/ enjoyment[4]	124	4.38	1.44	4.20	1.48	-0.17	1.2^{ns}
	Interest/ value or usefulness[5]	124	4.60	1.30	4.45	1.33	-0.15	1.2^{ns}
Finland	Perceived autonomy/choice[1]	27	4.61	1.15	4.22	1.10	-0.40	1.8^{ns}
	Perceived competence[2]	27	4.60	1.23	4.63	0.79	0.03	0.2^{ns}
	Support for relatedness[3]	27	4.70	1.09	4.49	0.86	-0.21	1.4^{ns}
	Interest/ enjoyment[4]	27	4.44	1.25	4.31	1.00	-0.13	0.7^{ns}
	Interest/ value or usefulness[5]	27	5.08	1.55	4.68	1.52	-0.40	1.9^{ns}
Greece (Florina)	Perceived autonomy/choice[1]	37	3.72	0.96	4.28	0.98	0.57	2.5*
	Perceived competence[2]	37	5.88	0.85	5.78	0.92	-0.10	0.6^{ns}
	Support for relatedness[3]	37	4.73	1.13	4.63	0.97	-0.10	0.5^{ns}
	Interest/ enjoyment[4]	37	5.89	1.06	5.79	0.91	-0.10	0.5^{ns}
	Interest/ value or usefulness[5]	37	6.08	0.99	6.12	1.10	0.04	0.2^{ns}
Greece (Thessaloniki)	Perceived autonomy/choice[1]	44	4.38	1.16	4.31	1.15	-0.07	0.6^{ns}
	Perceived competence[2]	44	4.64	1.24	4.84	0.94	0.21	2.0^{ns}
	Support for relatedness[3]	44	4.66	1.10	4.57	1.10	-0.10	1.0^{ns}
	Interest/ enjoyment[4]	44	4.81	1.40	4.80	1.36	-0.01	0.1^{ns}
	Interest/ value or usefulness[5]	44	5.46	1.42	5.29	1.39	-0.17	2.1*
Italy	Perceived autonomy/choice[1]	27	5.73	0.65	4.71	1.11	-1.02	4.3***
	Perceived competence[2]	27	4.45	0.92	4.84	0.71	0.37	1.9^{ns}
	Support for relatedness[3]	27	4.87	0.66	4.83	1.04	-0.04	0.2^{ns}
	Interest/ enjoyment[4]	27	5.78	0.82	5.44	0.93	-0.34	1.6^{ns}
	Interest/ value or usefulness[5]	27	5.47	0.86	5.94	0.77	0.46	3.0**
Spain	Perceived autonomy/choice[1]	70	3.89	1.04	4.09	0.93	0.20	1.4^{ns}
	Perceived competence[2]	70	4.56	1.01	4.15	1.13	-0.41	3.5***
	Support for relatedness[3]	70	4.80	0.67	4.87	0.74	0.08	0.8^{ns}
	Interest/ enjoyment[4]	70	4.67	1.29	4.32	1.38	-0.36	1.8^{ns}
	Interest/ value or usefulness[5]	70	5.12	1.28	4.70	1.23	-0.42	2.8**

*) ^{ns} $p > 0.05$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$; Examples of items in each subscale: [1]I do the activity because I want to do it; [2]I think I am pretty good at the activity; [3]I feel close to my peers during the activity; [4]I enjoy the activity very much; [5]I think doing the activity could help me to learn science

References

Deci, E. L., Eghrari, H., Patrick, B. C., & Leone, D. (1994). Facilitating internalization: The self-determination theory perspective. *Journal of Personality*, *62*, 119-142.

McAuley, E., Duncan, T., & Tammen, V. V. (1987). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly for Exercise and Sport, 60,* 48-58.

OECD (2007). *PISA 2006: Science Competencies for Tomorrow's World, Volume 1: Analysis.* Paris: OECD.

Osborne, J., Simon, S. & Collins, S. (2003). Attitude towards science: a review of the literature and its implications. *International Journal of Science Education, 25*(9), 1049–1079.

THE IN/CONSISTENCY OF STUDENTS' IDEAS ACROSS EVAPORATION, CONDENSATION, AND BOILING

Zubeyde Demet Kirbulut
Middle East Technical University

Michael Edward Beeth
University of Wisconsin Oshkosh

Abstract

A great amount of effort has gone into explaining the nature of students' ideas and the role that teachers and curriculum developers should play in addressing those ideas through instruction. Although there is no consensus on how best to address students' ideas during instruction, some researchers advocate that these ideas are coherent, systematic, and even theory-like while others describe them as incoherent, and fragmented. The purpose of this study was not to enter into the debate over the nature of students' ideas but rather to investigate an emerging issue - the consistency or inconsistency of features that should be expected across multiple related ideas in chemistry and the everyday phenomena that invoke these ideas. We were interested in students' abilities to use their ideas about evaporation, condensation, and boiling consistently when exposed to different representations of these concepts and to everyday contexts. We investigated the degree of consistency or inconsistency 12 high school students displayed when explaining multiple related ideas to understand what needs to be done to help students understand conceptually related ideas.

Introduction

Over the last three decades, a great amount of effort has been put on to explain the nature of naïve ideas. However, there is no consensus on the nature of naïve ideas among researchers. Some researchers advocated that naïve ideas are coherent, systematic, and even theory-like (e.g., Gopnik & Wellman, 1994; Vosniadou, 1992; Vosniadou, Vamvakoussi & Skopeliti, 2008), while the other researchers described naïve ideas as incoherent, fragmented, and not theory-like (e.g., diSessa, 1988, 2008; diSessa, Gillespie & Esterly, 2004). The purpose of this study is to investigate the consistency of students' ideas about evaporation, condensation, and boiling across representational, conceptual framework, and contextual consistency aspects defined in our data analysis. Also, we should note that we are neither on the side of coherence advocates nor on the side of fragmentation advocates. We believe that students can show a high degree of consistency or inconsistency in their naïve ideas. Furthermore, the interviews that we conducted with 12 high school students showed that some students had a high degree of consistency or inconsistency in their naïve ideas.

Rationale

Conceptual change research has been a paramount area in science education over the past three decades based on the foundations of constructivist learning. During this time, several conceptual change perspectives have been proposed (e.g., Chi & Slotta, 1993; diSessa, 1988; Gopnik & Wellman, 1994; Hewson, Beeth & Thorley, 1998; Hewson & Lemberger, 2000; Linder, 1993; Posner, Strike, Hewson & Gertzog, 1982; Vosniadou, 1992). These conceptual change perspectives differ in terms of the researchers' assumptions about the nature of learners' naïve ideas and the roles these ideas play in terms of future learning. They also differ in the extent to which they advocate for the coherence of naive ideas. We describe some researchers as advocates of extreme theory-like naïve ideas (e.g, Gopnik & Wellman, 1994) while others are advocates of knowledge in pieces (e.g., diSessa, 1988). There are other researchers who take positions in between these two extremes (e.g., Vosniadou, 1992). Likewise, coherence advocates who view naïve ideas in the light of scientific theory change are described in the literature as extreme. Less extreme coherence advocates would assert that naïve ideas are different from professional science by emphasizing a lesser degree of coherence between students' ideas and scientific ideas.

There has been wide agreement that understanding the nature of naïve ideas is important in science learning. However, beyond this agreement, researchers documented conflicting results germane to students' consistency of naïve ideas. While some researchers have proposed that students' naïve ideas were incoherent, fragmented, and not theory-like (e.g., BouJaoude, 1991; Cooke & Bredin, 1994; Engel Clough & Driver, 1986), others reported that students' naïve ideas were coherent and systematic (e.g., Vosniadou, 1992; Watson, Prieto & Dillon, 1997). Engel Clough and Driver (1986) explored the consistency of students' conceptions in pressure, heat, and biological evolution across a number of tasks. They found that students tended to use their alternative ideas inconsistently. Similarly, BouJaoude (1991) examined students' understanding about the concept of burning. He claimed that students' understanding about burning was fragmented, task specific, and context dependent. However, in their study of students' ideas about combustion, Watson et al. (1997) found a high degree of consistency in students' explanations about combustion. By contrast, in the field of force and motion, Cooke and Breedin (1994) reported that students' impetus theory explanations were highly inconsistent. Vosniadou (e.g., Vosniadou, 1992) studied students' conceptual knowledge in the domain of astronomy. Her conclusions stand on the "coherence" side of the debate (Vosniadou, 1992). She found that students tended to use both scientific and alternative conceptions coherently.

The debate concerning whether naïve ideas are coherent, systematic, and theory-like; or incoherent, fragmented, and not theory-like continues in conceptual change arena. In addition, many researchers involving in this debate are reporting conflicting results (e.g., diSessa et al. 2004; Vosniadou et al., 2008). It is crucial to elaborate the nature of naïve ideas to understand the process of conceptual change and then derive implications for instructional strategies that can impact these ideas. In this study, we will take a closer look at a range of students' the consistency of students' ideas about evaporation, condensation, and boiling across representational, conceptual framework, and contextual consistency aspects.

Evaporation, condensation, and boiling are fundamental concepts in most Chemistry curricula and courses. These concepts are also useful when explaining everyday phenomena. Many studies investigated students' conceptions related to these concepts (e.g., Canpolat, 2006; Gopal, Kleinsmidt & Case, 2004; Osborne & Cosgrove, 1983). Osborne and Cosgrove (1983) conducted clinical interviews with children from eight to 17 years of age to investigate their conceptions of the changes of states of water. They reported that children had superficial understanding about evaporation, condensation, boiling, and melting and those older children held similar views to younger children although they were exposed to considerable science teaching related to these concepts. Also, Gopal et al. (2004) interviewed second-year chemical engineering students and concluded that these students had inadequate understanding of evaporation and condensation. Similarly, Canpolat (2006) explored undergraduate students' misconceptions related to evaporation, evaporation rate, and vapor pressure. He found that students had superficial understanding related to these concepts, with the following main misconceptions related to evaporation: i) in order for evaporation to take place, a liquid has to take heat from its environment; ii) the evaporation rate of a liquid in an open container is different from that of the liquid in a closed container; iii) in a closed container, the evaporation rate decreases as time passes; and iv) the evaporation rate changes with surface area. Although there are many studies related to students' conceptions of evaporation, condensation, and boiling, there are no studies that investigate students' consistency of naïve ideas related to these three concepts collectively. Furthermore, understanding the nature of students' naïve ideas is crucial in making recommendations for science teaching and learning. However, in the literature, there are contradictory results concerning the nature of naïve ideas whether they are coherent, systematic, or even theory-like; or they are fragmented and incoherent. Moreover, there are few studies that could confirm the role of consistency across what are known to be naïve ideas. There is need for more in-depth studies to elaborate on the phenomenon of naïve ideas. Therefore, the purpose of this study is to examine the consistency of students' naïve ideas about evaporation, condensation, and boiling across representational, conceptual framework, and contextual consistency aspects in order to offer practical suggestions to classroom teachers.

Methods

This study examined the consistency of students' naïve ideas about evaporation, condensation, and boiling across representational, conceptual framework, and contextual consistency aspects. We chose the phenomenological method involving a small number of subjects (N=12) through extensive and prolonged engagement to develop patterns and relationships of meaning (Creswell, 1994). Data collection included semi-structured interviewing high school Chemistry students for roughly 55 minutes each.

Participants

Twelve students (seven females and five males) enrolled in a Midwestern high school in the USA consented to be interviewed for this study. Six students were enrolled in an advanced Chemistry course (ages 16 to 18) and six were enrolled in an introductory Chemistry course (ages 15 to 16). Purposeful typical case sampling method was used to identify students who were the focus of the study (Patton, 1990) with classroom teachers acting as key informants. In the findings for this study, alternatively, we present excerpts from interviews with two students: Linda, who exemplifies a high degree of consistency in her naïve ideas and Mark, who is an example of a high degree of inconsistency in his naïve ideas.

Procedure

The interview protocol was piloted and revised for face validity. During the interview, students were posed questions to examine their representational, conceptual framework, and contextual consistency across evaporation, condensation, and boiling. Interviews lasted up to 55 minutes, were video recorded and transcribed for analysis. The interview consisted of six questions and follow-up probes to investigate high school chemistry students' understanding of states of matter, melting, evaporation, condensation, boiling, and vapor pressure. For this study, we are presenting data related only to evaporation and condensation because of the space limitations. Table 1 shows sample questions for representational and contextual consistency about evaporation. The students were not asked specific questions for conceptual framework consistency. Their conceptual framework consistency was determined by examining their application of these concepts in a variety of contexts and in relation to other concepts.

Table 1. Sample questions for conceptual consistency about evaporation.

Representational Questions	Contextual Questions
Imagine an amount of ice in a closed glass cup at -25 ^{0}C (-13 ^{0}F). If the ice was heated to 125 ^{0}C (257 ^{0}F), what would happen to ice? Can you show where freezing, melting, evaporating, condensing and boiling occur on a graph?	In a room, there is an open plastic bottle half-filled with water. If this bottle were left for several days, what would happen to the level of water in the bottle?
How would you define evaporation?	After you wash your laundry and leave it for drying, what happens to water?
How can you draw the picture of evaporation in terms of the particulate nature of matter?	When pure water in an open container at 25 ^{0}C (77 ^{0}F) is left out to 10 ^{0}C (50 ^{0}F) for a while, what would happen to the level of water?

Data Analysis

Interview data were initially analyzed based on Creswell's (1994) six generic steps: i) organize and prepare the data for analysis, ii) read through all the data, iii) code the data, iv) generate themes or categories using the coding, v) organization and the description of the data in terms of the coding and themes, vi) interpretation of the data. The authors and their colleague independently coded the data, discussed the conflicts between categories, and the categories were finally verified. When we were examining the students' conceptual consistency, we used Savinainen and Viiri's (2008) characterization for conceptual coherence divided into three aspects: representational coherence (multiple representations such as drawings, graphics, and verbal representations), conceptual framework coherence (integration and differentiation between related concepts), and contextual coherence (application of a concept across a variety of situations). It should be noted that we used "consistency" instead of "coherence" since "consistency" refers to a student's answers in interviews while "coherence" refers to a student's internal mental models which were not the focus of this study.

Results

In this section, conceptual consistency of students' naïve ideas about evaporation and condensation across representational, conceptual framework, and contextual consistency aspects are analyzed.

Case 1: Linda

Linda is 16 years old, in 10th grade and has not taken any chemistry course before this class. She is planning to take two chemistry courses before graduated from high school. Her overall Grade Point Average (GPA) is 3.50 and she expects to receive a grade of B in the Chemistry course that she is currently attending.

Ideas about evaporation

First, Linda's multiple representations related to evaporation are presented to form the basis for her representational consistency. When asked how she would describe evaporation, she stated the following:

Evaporation is when a liquid changes phases to a gas, and once it reaches a certain temperature the liquid particles move faster because of the kinetic energy, and once it goes to a certain rate then the particles would break apart and be free moving and become a gas.

In her description of evaporation, Linda indicated that evaporation would occur at specific temperature when changing phase from liquid to gas. She also said that during evaporation, kinetic energy would increase and the particles would break apart. This statement is compatible with her drawing of the picture of evaporation of water in terms of particles and her explanation of where evaporation would occur on a phase change graph for water. In her drawing, she indicated that when evaporation occurred, kinetic energy increased and water molecules broke apart into hydrogen and oxygen gases. Furthermore, she identified the point at which evaporation of water occurred as 100 ^{0}C in her graph. Also, she verbally explained that below 100 ^{0}C, she did not think that any evaporation would occur.

Second, in terms of conceptual framework consistency, Linda demonstrated a high degree of consistency in her naïve ideas when relating evaporation to other concepts such as conservation of mass and chemical bonding. Her confusion when explaining conservation of mass related to evaporation can be seen in the following excerpt:

Researcher (R): If you would be able to measure the volume of water in a closed container, what would you expect that the volume of water to be after evaporating?
Linda (L): It would be the same because it doesn't lose any of the water.
R: And if you would be able to measure the weight of water in a closed container, what would you expect that the weight of water to be after evaporating?
L: It would be the same.

To this point, she did not integrate her understanding of chemical bonding with her knowledge of evaporation. She indicated that water particles would break apart into hydrogen and oxygen gases in her verbal explanation and her drawing of evaporation. She never considered that molecules were drawn together by intermolecular forces such as hydrogen bonding and when water evaporated, kinetic energy increased and hydrogen bonding between water molecules became weaker.

Third, she was consistent when expressing a naïve idea- the need for a specific temperature for evaporation to occur, even when this question was posed in different contexts. In the statement below, when asked if an open plastic bottle half-filled with water were left for several days, what would happen to the level of water in the bottle, she said that it would stay the same:

> R: There is an open plastic bottle half-filled with water. If this bottle were left for several days, what would happen to the level of water in the bottle?
> L: It would stay the same because there is no phase change at 25 ^{0}C.
> R: So at what temperature do you think the level of water would decrease?
> L: At 100 ^{0}C, it would start to evaporate.

When she was asked about an everyday phenomenon, wet clothes drying on a clothesline, she indicated the same naïve idea as above- that evaporation of water would only occur at 100 ^{0}C.

As all of the evidence above shows, Linda had conceptual consistency in her naïve ideas of evaporation across representational, conceptual framework, and contextual consistency aspects.

Ideas about condensation

Similar to her naïve ideas on evaporation, Linda showed a high degree of consistency in her naïve ideas related to condensation. In her verbal description of condensation, she indicated that condensation would occur whenever a gas turned into liquid at 100 ^{0}C. Her graphical representation of condensation is also compatible with this verbal explanation. She represented water condensing at 100 ^{0}C on the phase change graph and when asked to explain her graphical representation, she said that condensation would only occur at 100 ^{0}C:

> R: When does condensation occur?
> L: Condensation occurs toward end of this (along the horizontal liquid line she drew from right to left)
> R: Do you think in this part (along the vertical liquid line) there is any condensing?
> L: I don't think there is any condensing happening here.

Her drawing of the picture of condensation in terms of particles is consistent with her naïve idea about evaporation. She thought that during the evaporation of water, the particles would break apart into hydrogen and oxygen gases. To be consistent with this naïve idea, she also thought that when condensation of water occurred, hydrogen and oxygen gases would come together to form water molecules at 100 ^{0}C.

Furthermore, when asked to draw a picture of condensation related to everyday phenomenon which is asking the reason for water droplets formed on the outer surface of a bottle of cold liquid beverage, she showed water particles forming from hydrogen and oxygen gases in the atmosphere. This explanation is also consistent with her verbal explanation of this phenomenon. However, she did not demonstrate her naïve idea that the temperature must be at 100 ^{0}C for condensation to occur:

> R: Think about a bottle of liquid beverage which is taken out of a cold refrigerator. After some time, you see water droplets formed on the outer surface of the bottle. Where do you think the droplets come from?
> L: I am not sure from where. I am guessing it would be from air maybe because there is hydrogen in the air and there is oxygen in the air.
> R: So you mean that oxygen and hydrogen come together?
> L: Yes.

Second, in terms of her conceptual framework consistency, Linda demonstrated consistency in her naïve ideas when relating condensation to chemical bonding. She did not integrate her understanding of chemical bonding with her knowledge of condensation as thoroughly as she did for evaporation. Both her verbal explanation and drawing of condensation of water showed that hydrogen and oxygen gases would come together to form water.

Third, it was seen that she was consistent in her naïve ideas about condensation when she was posed questions about these concepts in different contexts. For example, she demonstrated her naïve idea that during condensation, hydrogen and oxygen gases would come together to form water. However, she did not demonstrate her naïve idea that the temperature must be at 100 ^{0}C for condensation of water as was seen above everyday phenomenon which is asking the reason for water droplets formed on the outer surface of a bottle of cold liquid beverage.

She also demonstrated the same naïve idea that during condensation, hydrogen and oxygen gases would come together to form water when water vapor condenses:

> R: When you hold your hand above boiling water, your hand gets wet. How can you explain this?
> L: Water is condensing because it is going from the gas state to the liquid state because your hand is colder than 100 ^{0}C. So whenever the gas hits your hand, it gets colder and becomes a liquid. But since it is pretty much directly from 100 ^{0}C to like maybe 99 ^{0}C, it's going to be really hot for your hand.
> R: What are the particles?
> L: Hydrogen and oxygen particles.
> R: And what is the wetness on your hand?
> L: That's water.

Moreover, the following excerpt shows that she had confusion regarding the temperature of condensation of water in a closed system. She could not consider saturated vapor concept to explain condensation of water:

> R: At room temperature, there is a tightly capped plastic bottle half-filled with water. If this bottle is left for several days, you can see many tiny water droplets appear on the lid of the bottle. Where do these water droplets come from?
> L: I think water inside is evaporating and since it's a closed container, they don't escape so it becomes a liquid again.
> R: But this system is at room temperature….
> L: Hmm. Then, I am not sure.

It is clear that Linda displayed conceptual consistency in her naïve ideas about condensation across representational, conceptual framework, and contextual consistency aspects.

Case 2: Mark

Mark, 16 years old and in 10th grade, has not taken any chemistry course before this class. He is planning to take three chemistry courses before graduating from high school. His overall GPA is 3.70 and he expects to receive a grade of A in the Chemistry course that he is currently attending.

Ideas about evaporation

First, in terms of representational consistency, Mark showed a high degree of inconsistency in his naïve ideas. When he was asked how he would describe evaporation, he defined evaporation as the release of the molecules from liquid phase to gas phase and he stressed the role of intermolecular forces at evaporation:

> Evaporation I guess it would have to be water changing into the gas form I guess… I guess evaporation would be the release of the molecules from the liquid phase to the gas phase… The molecules would begin to break apart [spread apart] so it would go so there are separate molecules and they're not bonded anymore and they just begin to release from the bottle just move up yea evaporate out I guess.

He also said that evaporation would occur at a specific temperature which was 100 ^{0}C for water, "…after 100 ^{0}C it would be the evaporation point….above 100 ^{0}C to 125 ^{0}C it would change to the gas phase".

He demonstrated his naïve idea that the temperature must be at 100 ^{0}C for evaporation to occur when he was drawing the phase change graph for water. He was also asked to draw a representation of evaporation in terms of particles. Although he indicated the role of intermolecular forces correctly in his one of drawings and in his verbal description, he thought that heat caused the evaporation within the molecule itself in his other drawing.

Second, in terms of conceptual framework consistency, he demonstrated inconsistency in his naïve idea about chemical bonding when relating chemical bonding to evaporation. As mentioned for his representational consistency, he first related his understanding of chemical bonding correctly to the evaporation concept; however, as he was posed other questions related to evaporation, such as his drawing of evaporation of water, he thought that hydrogen bonding occurred within individual molecules.

Third, he also showed a high degree of inconsistency in the application of his evaporation concept through different contexts. He showed inconsistencies when answering the same question as indicated below:

R: In a room, there is an open plastic bottle half-filled with water. If this bottle were left for several days, what would happen to the level of water in the bottle?
Mark (M): I would expect would [the level to] stay the same because no heat is added because the water stays the same at room temperature. So it wouldn't reach the evaporation point where water would evaporate from the bottle.
R: What do you mean by the evaporation point?
M: It wouldn't get to a certain temperature. I guess the heat around it wouldn't increase; the room temperature wouldn't increase so the point where the water would say get too hot to boil wouldn't go to like a boiling point. I would say it just stays the same because the room temperature would be constant.

The above excerpt also showed that Mark had confusion with his conceptions of heat and temperature and kinetic molecular theory. He did not consider that evaporation of water occurs at every temperature without heating by using its internal energy. Also, when he was describing evaporation, he changed his answer to the question immediately above:

R: You said you changed your answer. Why did you change your answer?
M: Well I think that because I would think that after so many days with it sitting that water would maybe…. tiny amounts evaporate from it. And also with the laundry question, that it will have to go to somewhere so I would think that the temperature in the air would be enough to evaporate it even though it's very small amounts.

Furthermore, the following excerpt shows his application of evaporation concept in an everyday phenomenon, wet clothes drying on a clothesline:

R: After you wash your laundry and leave it for drying, what happens to the water?
M: The water would evaporate off, I guess.
However, he could not apply this evaporation concept to the following context:
R: When pure water in an open container at 25 ^{0}C is left out to 10 ^{0}C, what would happen to the level of water?
M: The level of water I think would stay the same. So the temperature is decreasing obviously so yea I think it would be the same because it's not to that freezing point where it becomes a solid or anything like that so I'm pretty sure that it stays the same because none would evaporate out because the temperature is being lowered.

As all of the evidence above shows, Mark had conceptual inconsistency in his naive ideas about evaporation across representational, conceptual framework, and contextual consistency aspects.

Ideas about condensation

As he did for evaporation, Mark showed a high degree of inconsistency in his representation of condensation. He defined condensation as follows:

> Condensation is going from the gas phase back into the liquid phase so it's very hot temperature and then it's cooled back down that evaporated water molecules may turn back to liquid molecules may slow down, come together… So it is say it's outside the cup it would evaporate out and the temperature drops. The water molecules may then again come back together and bond again after. So as the temperature drops these begin to slow down, all the molecules slow down and they can bond again to form hydrogen bonds to become liquid.

Although he defined condensation using intermolecular forces correctly, when he was asked to draw a picture of condensation of water in terms of particles, he could not represent condensation of water in terms of intermolecular forces. He thought that hydrogen bonding occurred within molecules as he did for evaporation. Also, it should be noted that he could not apply the concept of saturated vapor to explain condensation in a closed system. He thought that condensation only occurred because of temperature change.

Second, in terms of conceptual framework consistency, he demonstrated inconsistency in his naïve idea when relating condensation to chemical bonding as stated above. He also had confusion in relating saturated vapor and heat and temperature concepts to condensation. For example, he thought that there must be an abrupt change in temperature for condensation to occur.

Third, he demonstrated a high degree of inconsistency in his naive ideas about condensation through different contexts. For example, when he was asked his idea about the water droplets formed on the lid of the closed bottle half-filled with water and left for several days, he could not apply his understanding of chemical bonding thoroughly to explain condensation of water. Furthermore, he did not consider saturated vapor as a concept when explaining that phenomenon:

> I think it is coming from the evaporation. It's trying to get release from the bottom into the liquid phase but since there is a cap on it gets stopped there where they again bond back to the liquid form. When it's liquid in the bottom again the molecules would break apart, turn into the gas phase and move up but again they stay bonded once it again turns into the liquid phase.

It should be noted that during the interview, Mark demonstrated somehow correct reasoning relating to chemical bonding to condensation and evaporation concepts, but as shown above, he could not apply that understanding consistently for all of the questions, even though he indicated that he was aware of his inconsistency.

When he was asked another everyday phenomenon question related to the reason for water droplets forming on the outer surface of a bottle of cold liquid beverage, he could not explain how condensation occurred. He thought that water droplets were caused by the temperature difference from outside and inside the bottle.

The following excerpt shows another example of his application of condensation in a different context. Although it was seen that his reasoning was somehow correct in terms of intermolecular forces, he again showed his naïve idea about evaporation - that evaporation occurred at a certain temperature:

> R: When you hold your hand above the boiling water, your hand gets wet. How can you explain this?
> M: I explain to where it's at the boiling point so there is evaporation so evaporating water moves up and it goes into your hand which isn't at the same temperature as boiling water so that may turn back into the liquid phase, they may bond back together.

Evidence presented above showed that Mark had conceptual inconsistency in his naïve ideas about condensation across representational, conceptual framework, and contextual consistency aspects.

Conclusions and Implications

In the literature, there is no consensus on whether students use their naïve ideas consistently or inconsistently. Some researchers advocated that naïve ideas are coherent, systematic, and even theory like (e.g., Gopnik & Wellman, 1994; Vosniadou, 1992, Vosniadou et al., 2008), while the other researchers described naïve ideas as incoherent, fragmented, and not theory-like (e.g., BouJaoude, 1991; diSessa et al., 2004; Engel Clough & Driver, 1986) As we indicated before, we stand neither on the side of coherence advocates nor on the side of fragmentation advocates. Indeed, the findings of this study indicated that students had a high degree of consistency and inconsistency in their naïve ideas.

This study showed that although students were taught evaporation, condensation, and boiling concepts from elementary school, they still do not have deep understanding of these concepts. Teaching and learning chemistry requires integration of three perspectives which are macroscopic, submicroscopic, and symbolic (Johnstone, 1993). Many students have difficulties in relating and making transitions among these three perspectives (De Jong & Taber, 2007). In this study, evidence from the interviews showed that the students could not make transitions among these perspectives.

The findings of this study together with contradictory research findings in the literature cause us to question how conceptual change can be promoted. Conceptual change perspectives differ in terms of researchers' assumptions about the nature of learners' naïve ideas. However, these conceptual change perspectives indicated the importance of metaconceptual awareness independent of their assumptions on naïve ideas (e.g., Hewson et al., 1998; Mortimer, 1995; Vosniadou et al., 2008). For example, Hewson et al. (1998) pointed out that metaconceptual activities were inherent in conceptual change. Similarly, Vosniadou et al. (2008) indicated that metaconceptual awareness was crucial for conceptual change to be achieved. The findings of this study highlighted the importance of metaconceptual teaching practices in facilitating students' conscious use of their ideas through different representations, contexts and concepts. For example, when Mark was asked the difference between evaporation and boiling, he said that the temperature must be 100 ^{0}C for boiling and at or above 100 ^{0}C for evaporation of water to occur. However, when he was answering an everyday phenomenon related to evaporation of water, he gave correct answer by saying that evaporation could also occur at room temperature. After he was asked for this inconsistency, he said that he was aware of his inconsistency about his responses related to evaporation and boiling; however, he did not know how to explain this. This finding suggests that teachers should take into account metaconceptual teaching practices in their classrooms.

Some practical implications from this study are that teachers should expect students to relate and make transitions among macroscopic, submicroscopic, and symbolic perspectives when teaching evaporation, condensation, and boiling. Students should be exposed to questions in which they need to use different representations, apply their ideas in different contexts, and integrate and differentiate among related concepts about evaporation, condensation, and boiling. Our results also suggests that to have deep understanding of evaporation, condensation, and boiling concepts, students not only need to differentiate contextually their ideas and representations but also they must metacognitively aware of their ideas and integrate among related concepts about evaporation, condensation, and boiling concepts. Therefore, teachers should be aware of the effect of metaconceptual teaching practices in facilitating conceptual change.

References

BouJaoude, S.B. (1991). A study of the nature of students' understandings about the concept of burning. *Journal of Research in Science Teaching, 28*(8), 689-704.

Canpolat, N., Pinarbasi, T., & Sozbilir, M. (2006). Prospective teachers' misconceptions of vaporization and vapor pressure. *Journal of Chemical Education, 83*(8), 1237-1242.

Chi, M.T.H., & Slotta, J.D. (1993). The ontological coherence of intuitive physics. Cognition *and Instruction, 10*(2&3), 249-260.

Cooke, N.J., & Bredin, S.D. (1994). Constructing naïve theories of motion on the fly. *Memory & Cognition, 22*(4), 474-493.

Creswell, J.W. (1994). *Research Design Qualitative and Quantitative Approaches.* Sage Publications, Inc.: Thousand Oaks, California.

De Jong, O., & Taber, K.S. (2007). Teaching and learning the many faces of chemistry. In S.K. Abell, & N.G. Lederman (Eds.), *Handbook of Research on Science Education* (pp. 631-652). Mahwah, NJ: Erlbaum.

diSessa, A.A. (1988). Knowledge in pieces. In G. Forman, & P. Pufall (Eds.), *Constructivism in the computer age* (pp. 49-70). Hillsdale, NJ: Erlbaum.

diSessa, A.A. (2008). A bird's-eye view of the "pieces" vs. "coherence" controversy (from the "pieces" side of the fence). In S. Vosniadou (Eds.), *International Handbook of Research on Conceptual Change* (pp. 35-60). Routledge: New York.

diSessa, A.A., Gillespie, N., & Esterly, J. (2004). Coherence vs. fragmentation in the development of the concept of force. *Cognitive Science, 28,* 843-900.

Engel Clough, E., & Driver, R. (1986). A study of consistency in the use of students' conceptual frameworks across different task contexts. *Science Education, 70*(4), 473-496.

Gopal, H., Kleinsmidt, J., & Case, J. (2004). An investigation of tertiary students' understanding of evaporation, condensation and vapor pressure. *International Journal of Science Education, 26*(13), 1597-1620.

Gopnik, A., & Wellman, H.M. (1994). The theory theory. In L.A. Hirschfeld & S.A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 257-293). New York: Cambridge University Press.

Hewson, P.W., Beeth, M.E., & Thorley, N.R. (1998) Teaching for conceptual change. In K.G. Tobin & B.J. Fraser (Eds.), *International Handbook of Science Education* (pp. 199-218). Dordrecht, Netherlands: Kluwer.

Hewson, P. W., & Lemberger, J. (2000). Status: The hallmark of conceptual learning. An example from learning genetics. In R. Millar, J. Leach, & J. Osborne (Eds.), *Improving science education* (pp. 110-125). London: Open University Press.

Johnstone, A.H. (1993). The development of chemistry teaching. *Journal of Chemical Education, 70*(9), 701-705.

Linder, C.J. (1993). A challenge to conceptual change. *Science Education, 77* (3), 293-300.

Mortimer, E.F. (1995). Conceptual change or conceptual profile change? *Science & Education, 4,* 267-285.

Osborne, R.J., & Cosgrove, M.M. (1983). Children's conceptions of the changes of state of water. *Journal of Research in Science Teaching, 20*(9), 825-838.

Patton, M.Q. (1990). *Qualitative Evaluation and Research Methods.* Sage Publications, Inc.: Newbury Park.

Posner, G.J., Strike, K.A., Hewson, P.W., & Gertzog, W.A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, *66*, 211-227.

Savinainen, A., & Viiri, J. (2008). The force concept inventory as a measure of students' conceptual coherence. *International Journal of Science and Mathematics Education, 6,* 719-740.

Vosniadou, S. (1992). Knowledge acquisition and conceptual change. *Applied Psychology: An International Review, 41*(4), 347-357.

Vosniadou, S., Vamvakoussi, X., & Skopeliti, I. (2008). The framework theory approach to the problem of conceptual change. In S. Vosniadou (Eds.), *International Handbook of Research on Conceptual Change* (pp. 3-34). Routledge: New York.

Watson, J.R., Prieto, T., & Dillon, J.S. (1997). Consistency of students' explanations about combustion. *Science Education, 81,* 425–444.

THE ROLE OF MATHEMATICS IN PHYSICS – THE STUDENTS' POINT OF VIEW

Olaf Krey
University of Potsdam, Germany

Helmut F. Mikelskis
University of Potsdam, Germany

Abstract

There are several empirical studies (qualitative as well as quantitative research) trying to describe students' mathematical world views as well as their views about the nature of science (e.g. physics) and the nature of scientific knowledge. Unquestionable, mathematics is essential for the physics we do and teach at our schools and universities. Surprisingly enough, there is no research available that one could use to answer the question about students' perspectives on the role of mathematics in physics. The project upon which this paper is based tries to close this gap by providing a first inventory on students' views about the role of mathematics in physics. In the following, examples from this inventory are shortly presented in order to give an overview of concepts, methods and strategies we applied to produce empirical findings that serve as a base for further research.

Introduction: Background and objectives

Attitudes, beliefs and belief systems, subjective theories etc. are a well established field of research in educational psychology, in mathematics education (mathematical world views) and in science education (images of science). Important publications include (Carey/Smith 1993), (Grigutsch 1996), (Hammer 1994), (Hammer 1995), (Hofer/Pintrich 1997), (Köller/Baumert/Neubrand 2000), (Meichtry 1992), (Meichtry 1993), (Ryder/Leach/Driver 1999), (Schoenfeld 1983), (Schoenfeld 1992), (Törner/Grigutsch 1994).

The motivation for this kind of research is the underlying assumption, that all these intuitive theories influence our approaches towards the world around us, that they influence the way we perceive things, process information, think and learn. This instrumental reason for studying beliefs is complemented by a substantial one – namely, these beliefs are core components of what we call mathematical or scientific literacy (cf (Carey/Smith 1993), (Meichtry 1993), (Labudde 1998), (Schoenfeld 1992)). From a physics education point of view one can say, that all reasons one can find to conduct research on students' conceptions about the nature of science (nos) ((Driver et al. 1996)) also suggest research on the role of mathematics in physics, since it is an essential component of the nature of physics and is involved in almost all activities. While there are a few philosophical approaches towards exploring the role of mathematics in physics, this field is not covered by empirical research on students beliefs.

In our study we have developed a questionnaire, that was given to physicists, in service physics teachers, physics teacher students and pupils grades 10 and 12. In this paper we focus on the physics teacher students. In the following we exemplify that physics teacher students' conceptions about the role of mathematics in physics differ from those held by pupils grade 10, that constructs from research about NoS can be successfully applied in our field of interest and that one can identify certain types of teacher students.

The popularity of mathematical activities in physics

Physics is a very unpopular school subject to (not only) German pupils (Gardner 1987). Especially the interest „to calculate something“ was proven to be very low (Hoffmann et al. 1998). Our interest lies in exploring in more detail, what exactly students do not like about the mathematical activities (MA) they are confronted with. The results of a pilot study (N=282, German Gymnasium pupils grade 10; cf Krey 2008) will be compared to data gathered from German physics teacher students (N=119, 3rd or 5th semester).

Students were asked to rank the following 8 mathematical activities (cf Liebers 1983): defining physical quantities by use of equations (MA 1), recording measured values in tables and their graphical representation (MA 2), evaluating measured values by use of equations (MA 3), setting up equations to describe physical assumptions and conditions (MA 4), deduction of physical laws (MA 5), explaining or predicting physical processes by interpretation of equations (MA 6), explaining or predicting physical processes by interpretation of diagrams (MA 7), calculation of physical quantities (MA 8). Results are shown in figure 1.

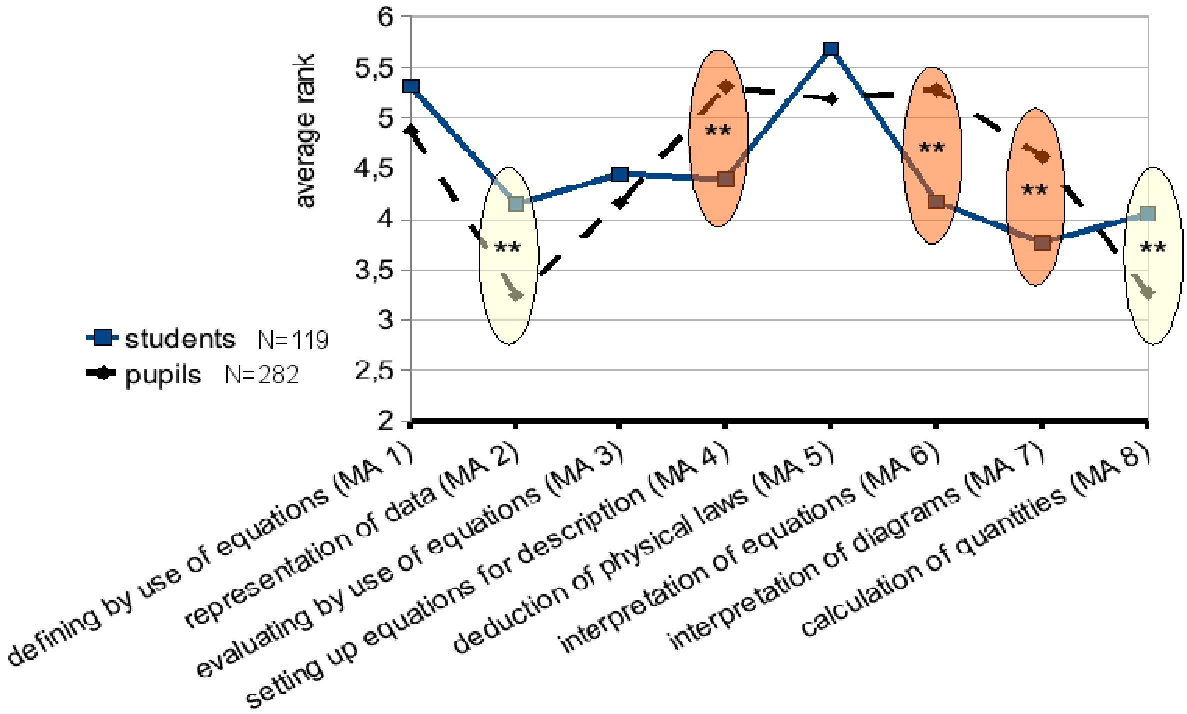


Figure 1. Students' and pupils' ranking on the popularity of mathematical activities

For grade 10 pupils less demanding cognitive tasks (rote processing) are more popular to pupils than those that demand higher skills needed for modelling tasks, e.g. MA 2 and MA 8 are the most popular activities ($p<.01$). Reasons are most likely to be found in instruction design, tendency to avoid cognitive work load and better payoff from those activities. For more details see Krey 2008. Grade 10 pupils and 3rd semester physics teacher students differ in their judgements about these activities. (Statistical analysis – assuming ordinal (Friedman-Test, Wilcoxon-Test) or metric (t-test) data – identify MA 5 as the least popular activity ($p<.01$) for students, while MA 6 and 7 are the most popular ones ($p<.01$).)

It is promising, that students rank formal activities as MA 2, 8 as significantly less popular ($p<.01$) than grade 10 pupils and at the same time rate modelling related activities such as MA 4, 6, 7 as significantly more popular ($p<.01$) as shown in figure 1. The impression of judgement based on cognitive workload found for the pupils cannot be seen here. Of course, one needs to keep in mind, that we asked physics teacher students to rank the activities. In general they are the former most successful pupils. However, statistical analysis shows that high achieving pupils' rankings do (in average) not differ significantly from those of low achieving pupils. This seems to indicate a change in interest or beliefs, but also needs more research to be confirmed.

Proximal and distal conceptions about the role of mathematics in physics

Beliefs concerning different aspects of nos have been analyzed in the context of mathematics and physics education. In fact, contradicting results have been found. Hogan holds the opinion that the distinction between distal conceptions (beliefs and knowledge about the scientific enterprise) and proximal conceptions (beliefs and knowledge about nos as experienced in the pupils' school context) might help to explain those inconsistencies (cf Hogan 2000).

If Hogan is right, then this would not only show practical relevance for teaching about the nature of science, but also have implications for future nos research and theory. There also are follow up questions that need to be addressed. Is this distinction a general necessity or could it be dependent on the content? Drawing on our experience from other fields of research we can expect that this is most likely the case.

In this study, we tried to find empirical evidence for Hogan's theory based distinction between proximal and distal conceptions. A result concerning the role of mathematics as a communication tool will be presented. Theoretical considerations suggest that mathematical elements such as equations and diagrams serve as communication tools (cf Frey 1967). Data to evaluate students' subjective perception of this communication function (N=130) was acquired, using a five-point Likert scale, identified by explorative factor analysis and analysis of reliability (see table 1). Mean values, significant results of t-tests and effect strengths are shown in figure 2, where 'Comm' stands for communication, 'D' for diagram, 'E' for equation, 'p' for proximal and 'd' for distal.

Table 1. Perception of communication function (CF) of diagrams and equations (proximal und distal)

	proximal conceptions	distal conceptions
CF of diagrams	Comm-D-p: 5 items; $\alpha_{Cronbach}=0,80$ Item example: It is easier for me to explain something in physics, if I can use diagrams.	Comm-D-d: 5 items; $\alpha_{Cronbach}=0,73$ Item example: In physics as a science, it is easier for physicists to explain something, if they can use diagrams.
CF of equations	Comm-E-p: 5 items; $\alpha_{Cronbach}=0,85$ Item example: It is easier for me to explain something in physics, if I can use equations.	Comm-E-d: 5 Items; $\alpha_{Cronbach}=0,70$ Item example: In physics as a science, it is easier for physicists to explain something, if they can use equations.

On the one hand data support Hogan's suggestion to clearly distinguish between proximal and distal conceptions (significant difference ($p<.01$) between distal and proximal subjective perception of communication function using equations). On the other hand it seems to be content-specific, whether this distinction is relevant or not (no significant differences between proximal and distal subjective perception of communication function using diagrams). Learners tend to perceive mathematical elements differently – perception of diagrams and equations is always (not only in this example) in favor of the first, which corresponds to our expectations. Together with the next example result of our research this suggests, that diagrams should be used more effectively in instructional settings, especially with low achieving students.

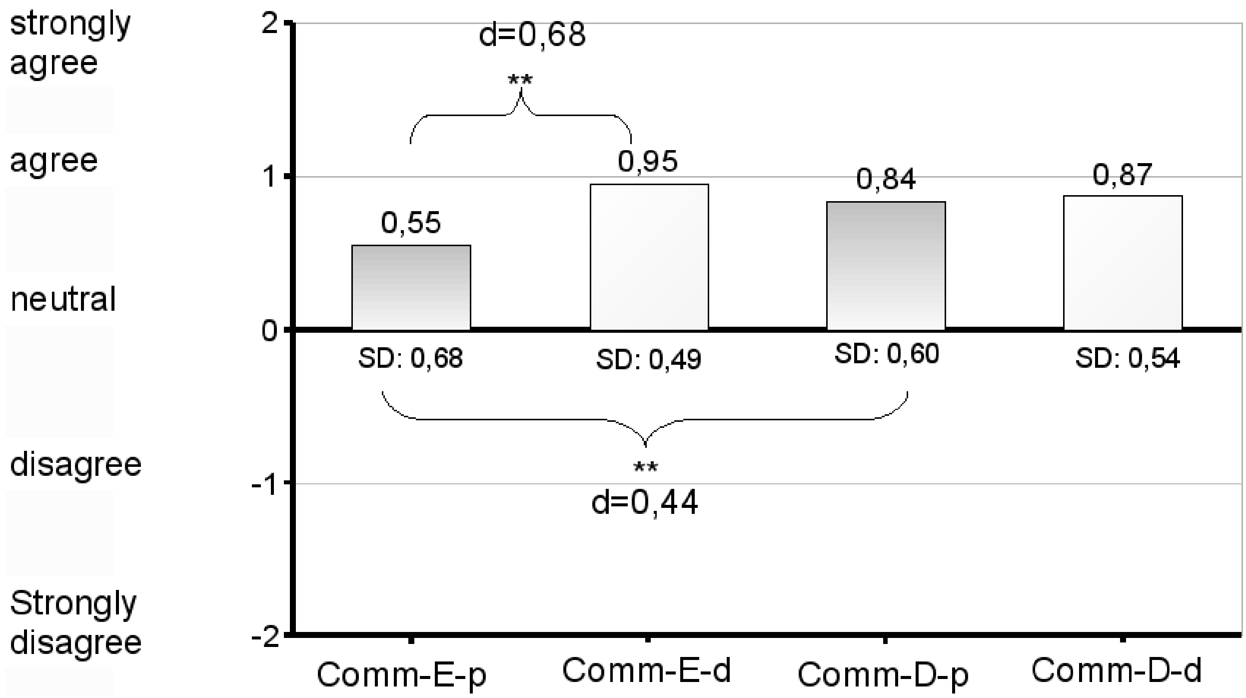


Figure 2. Subjective perception of communication function of diagrams and equations (proximal and distal)

Physics teacher students' performance and self-feeling when using mathematics in physics

In order to describe students' conceptions about the role of mathematics and physics, we were trying to identify „types“. In this case we applied a Latent-Class-Analysis (LCA) to the following data: sum score of a test, in which processes from reality presented in text had to be translated into equations and diagrams, the interest in theoretical physics, as well as the self-feeling when dealing with diagrams and equations. We found a 3-class solution (based on BIC, CAIC and bootstrapping) as shown in figure 3.

Classes 1 and 3 are easy to interpret (high-achieving and interested (class 3) and low-achieving and not interested (class 1). However, in class 1 dealing with diagrams is favoured over equations ($p<.05$). From this point of view, it is plausible to assume that physics instruction should be based on diagrams for low-achieving students. However, to confirm this assumption, more research is needed.

Class 2 students show high interest in theoretical physics, perform average and specify a preference of equations over diagrams. Whether this can be interpreted as a tendency to increase the feeling of self-worth by pretending to be able to understand abstract thoughts in abstract representational systems needs to be discussed. If this is the case, one can assume, that such attitudes will influence present learning and future teaching activities of these teacher students. It also shows, that one should think about how appropriate those conceptions and beliefs are and how one can address them more effectively in our physics teacher courses.

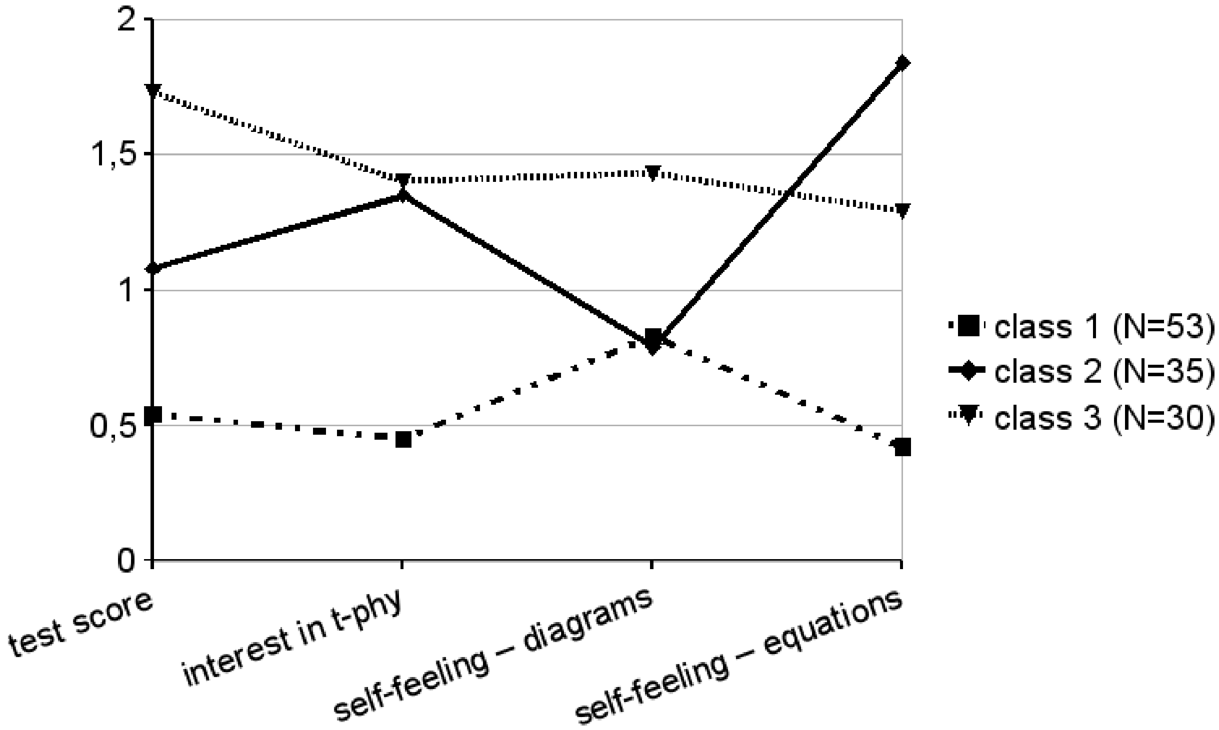


Figure 3. 3-class solution of LCA

Conclusion

In summary our results clearly suggest more detailed research on certain aspects of students' conceptions about the role of mathematics in physics. Data provide reasons to believe that those conceptions develop over the years, which should be investigated in a longitudinal study. Concepts taken from nos research have proven to be useful in our context, which provides strong arguments to include the role of mathematics as a core element of our students' conceptions about the nature of science into future research activities in this field. Finally we were able to identify different classes of students who clearly differ in their perception of the role of mathematics in physics. One can assume that these differences do affect both, the learning and later teaching of physics content of these teacher students.

References

Carey, S. & Smith, C. (1993). On understanding the nature of scientific knowledge. *Educational Psychologist*, *28*, 235-251.
Carey, S. & Smith, C. (1993). On understanding the nature of scientific knowledge. *Educational Psychologist*, *28*, 235-251.
Driver, R., Leach, J., Millar, R. & Scott, P. (1996). Young People's Image of Science. Buckingham, Philadelphia.
Frey, G. (1967). Die Mathematisierung unserer Welt. Stuttgart.
Gardner, P. L. (1987). Schülerinteressen an Naturwissenschaften und Technik. In M. Lehrke & L. Hoffmann (Eds.), *Schülerinteressen am naturwissenschaftlichen Unterricht* (Volume , pp. 13-38). Köln.
Grigutsch, S. (1996). Mathematische Weltbilder von Schülern. Struktur, Entwicklung, Einflußfaktoren. , , .
Hammer, D. (1994). Epistemological beliefs in introductory physics. *Cognition and Instruction*, *12*, 151-183.
Hammer, D. (1995). Epistemological considerations in teaching introductory physics. *Science Education*, *74*, 393-413.

Hofer, B. K. & Pintrich, P.R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, *67*, 88-140.

Hoffmann, L., Häussler, P. & Lehrke, M. (1998). Die IPN-Interessenstudie. Kiel.

Hogan, K. (2000). Exploring a Process View of Students' Knowledge about the Nature of Science. *Science Education*, *84*, 51-70.

Köller, O., Baumert, J. & Neubrand, J. (2000). Epistemologische Überzeugungen und Fachverständnis im Mathematik- und Physikunterricht.. In J. Baumert, W. Bos & R. Lehmann (Eds.), *TIMSS/III. Dritte Internationale Mathematik- und Naturwissenschaftsstudie - Mathematische und naturwissenschaftliche Bildung am ende der Schullaufbahn. Band 2: Mathematische und physikalische Kompetenzen am Ende der gymnasialen Oberstufe.* (Volume 2, pp. 229-270). Opladen: Leske+Budrich.

Krey, O. & Mikelskis, H.F. (2008). Mathematik im Physikunterricht aus Sicht der Schülerinnen und Schüler. In D. Höttecke (Ed.), *Kompetenzen, Kompetenzmodelle, Kompetenzentwicklung. Gesellschaft für Didaktik der Chemie und Physik. Jahrestagung in Essen 2007* (Volume , pp. 158-160). Münster.

Labudde, P. (1998). Konstruktivismus im Physikunterricht der Sekundarstufe SII. Bern: Universität Bern.

Liebers, K. (1983). Anwendung der Mathematik im Physikunterricht. Berlin.

Meichtry, Y. J. (1992). Influencing student understanding of the nature of science: Data from a case of curriculum development. *Journal of Research in Science Teaching*, *28*, 389-407.

Meichtry, Y. J. (1993). The impact of science curricula on student views about the nature of science. *Journal of Research in Science Teaching*, *30*, 429-443.

Ryder, J., Leach, J. & Driver, R. (1999). Undergraduate Science students' image of science. *Journal of Research in Science Teaching*, *36*, 201-219.

Schoenfeld, A. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. In D. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (Volume , pp. 334-370). New York: Macmillan.

Schoenfeld, A. H. (1983). Beyond the Purely Cognitive: Belief Systems, Social Cognition, and Metacognitions as Driving Forces in Intellectual Performance. *Cognitive Science*, *7*, 329-363.

Törner, G. & Grigutsch, S. (1994). "Mathematische Weltbilder" bei Studienanfängern - Eine Erhebung. *Journal für Mathematik-Didaktik*, *15*, 211-251.

PREFERRED LEARNING STYLES OF STUDENTS IN A COLLEGE OF AGRICULTURE

Ahmad Rezvanfar
University of Tehran

Mahtab Pouratashi
University of Tehran

Abstract

The main purpose of this study was to determine the preferred learning styles of students in a college of agriculture. The population consisted of senior students in college of Agriculture at the University of Boo-Ali-Sina, Iran. A sample of 80 students was selected by using random sampling method. The Group Embedded Figures Test (GEFT), as a standardized test, was applied to assess the preferred learning style of agriculture students as either field-independent or field-dependent. The data were analyzed by the use of statistical methods such as frequency distribution, percentage, mean, standard deviation, correlation analysis, and t-test. Findings of the study showed that 27.3% of male students preferred field-dependent learning style, while 72.2% of those preferred field-independent learning style. The comparison of students' GEFT scores by gender indicated that the male students' GEFT mean score was significantly higher than the female students' GEFT mean score. Also, the correlation analysis indicated that there was no relationship between the GEFT scores and age of respondents. Scholastic success of students in relation to learning styles indicated that the grade point average (GPA) for students with field-independent learning style was significantly higher than the grade point average (GPA) for students with field-dependent learning style.

Keywords: *Learning Style, Field-independent, Field-dependent, Agriculture, Student.*

Introduction

Technology is changing more rapidly than ever before and society is requiring that the workforce continually gain new knowledge to remain productive (Weisburg & Ullmer, 1995). Therefore, it is clear that someone who has learned how to learn will be a productive workforce. But what is the learning?

Learning as Kolb (1984) defined is "the process whereby knowledge is created through the transformation of experience" and occurs "through the active extension and grounding of experience and ideas in the external world and through internal reflection about the attributes of these ideas and experiences."

During the process of learning many factors influence that one of them is learning style. Learning style can be described as stable characteristics of an individual (Garger and Guild, 1984), consisting of distinctive behaviors which serve as indicators of how person learns from and adapts to her/his environment (Gregorc, 1979).

In considering learning styles, one of the learning style models is field-independent and field-dependent. The influence the surrounding field has on individual's perception of items within the field as well as its impact on the individual's intellectual domains and personality traits have been studied (Witkin, 1973; Raven et al, 1993; Cano & Metzger, 1995; Shih & Gamon, 2001; Garton et al, 2002). Witkin (1973) has shown that a person who perceives items as more or less separate from the surrounding field is said to be leaning toward a field-independent learning style. Field-independent learners tend to view the world more analytically. They are more likely to favor inquire and independent study and rely on self-defied goals (Witkin et al, 1977; Garger & Guild, 1984).

In contrast, a person whose mode of perception is strongly dominated by the surrounding field is said to be leaning toward a field-dependent learning style. Field-dependent learners tend to perceive the world globally. They are likely to favor the spectator approach to learning and require externally defined goals and are socially oriented (Witkin et al, 1977; Garger & Guild, 1984).

Learning styles is a key factor in several areas such as education and performance of students (Witkin, 1973) and also, career planning. Education has become a commodity that people seek to invest for their own personal gain and as a route to a better life (Davies, 1998). Educational researchers have shown that students are unique in their own ways, including the way they learn (Witkin, 1973; Gregorc, 1979). So, one of the most significant challenges that university faculties face is to be perceptive enough to recognize learning differences among their students.

Also, learning style is important in career selection. Individuals preferring field-dependent learning style will tend to choose occupations that require involvement with others. While individuals preferring field-independent learning style will tend to favor occupations where there is less emphasis on interpersonal interaction (Dembo, 1988).

Since learning styles affect students' learning outcomes, there is the need to understand learning styles of students in order for colleges and individuals to benefit.

Whereas one of the factors in reaching to agriculture development is accessing to educated workforce and one of the main institutions for educate workforce in agriculture are colleges of Agriculture, in the present study the respondents are agriculture students.

According to importance of learning styles in the process of education especially in the agricultural higher education and consequences of applying learning style, this paper will first attempt to investigate preferred learning styles of students in a college of agriculture

The special objectives of this study were:

- To study frequency distribution of respondents in relation to gender and age
- To study frequency distribution of respondents according to their learning style by gender.
- To make comparison of students' learning styles in relation to gender,
- To investigate scholastic success of students in relation to learning styles.

Today's society is requiring that the workforce continually gain new knowledge to remain productive. So it is clear that someone who has learned how to learn will be a productive member of the workforce. Learning process occurs within the individual, and during this process numerous factors influence that one of them is learning style.

Learning style is an important factor in several areas such as move to improve curricula, how teachers teach and students learn, interaction between teachers and students, academic performance of students, and career planning. Knowledge of learning styles will help instructors to be more insightful about how to adapt instruction to students' learning styles. In addition, knowledge of students' learning styles is a key factor for college counselors to aid students in career planning. Dembo (1988) mentioned that students preferring field-independent learning style will tend to favor occupations where there is less emphasis on interpersonal interaction, while students preferring field-dependent learning style will tend to choose occupations that require involvement with others.

Different studies have focused on investigating the relationship between agricultural students' learning styles and performance in agriculture courses, academic performance as measured by grade point average, and overall success in higher education (Garton et al, 2002; Cano, 1999; Daley et al, 1997). These studies revealed that when learning styles were considered in teaching and learning process, students' performance was enhanced.

The results of this study will help Iranian universities better understand their agricultural students' learning styles. Prior to this study, Iranian agricultural students' learning styles had not been studied.

Methods

To study the Preferred learning styles of students in a college of agriculture a sample of 80 students were selected from 400 students studying in the semester 7 and 8 at B.Sc. level in the Faculty of Agriculture, University of Boo-Ali-Sina, Iran by the use of a "multi-level stratified random sampling" method. A pre-tested questionnaire was used to collect the dada.

For the purpose of the study, students' learning style was measured by use of Group Embedded Figures Test (GEFT) which developed by Witkin, et. al (1971). The GEFT is a standardized instrument and the division between field-independent and field-dependent was set at a score of 12, as recommended by Witkin et al (1971). Individuals scoring 12 or above on the GEFT were considered as field independent and individuals scoring 11 or below were considered as field dependent.

Possible scores on the GEFT ranged from 0 to 18. In this study, the division between field-independent and field-dependent was set at a score of 12, as recommended by Witkin, et. al (1971). Individuals scoring 12 or above on the GEFT were considered as field independent. Individuals scoring 11 or below were considered as field dependent. The GEFT is a standardized instrument and validity and reliability of the GEFT was established by the authors of the instrument. Reliability for the GEFT was .82 (Witkin, et. al, 1971). In addition, scholastic success of agriculture students was measured by grade point average (GPA) at the completion of the junior academic year.

Collected data was analyzed using the Statistical Package for the Social Sciences (SPSS). To reach the research objectives appropriate statistical procedures were used. Data analysis was carried out in two sections, consisting description and inferential analysis. Statistics such as frequencies, percentage, mean, and standard deviation were used in the descriptive section. T-test and correlation analysis were used in the inferential analysis section.

Results

Frequency distribution of respondents in relation to their gender and age

Frequency distribution of respondents in relation to gender and age is shown in Table 1. As it can be seen, about 55.0% of the respondents were male and 45.0% were female. Respondents were on average 23 years old and most of them were arranged between the age of 23 and 25 years.

Table 1. Frequency distribution of respondents in relation to gender and age

%	Frequency	
	Gender	
55.0	44	Male
45.0	36	Female
100.0	80	Total
	Age	
31.2	25	< 23
53.8	43	23-25
15.0	12	> 25
100.0	80	Total

Frequency distribution of respondents according to their learning style by gender

Frequency distribution of respondents according to their learning style by gender is shown in table 2. The results showed that 27.3% of male students preferred field-dependent learning style, while 72.2% of those preferred field-independent learning style. 58.3% of female students preferred field-dependent learning style and 41.7% of those preferred field-independent learning style.

Table 2. Frequency distribution of respondents according to their learning style by gender

Gender		Frequency	Percent
Male	Field-Dependent	12	27.3
	Field-Independent	32	72.7
	Total	44	100.0
Female	Field-Dependent	21	58.3
	Field-Independent	15	41.7
	Total	36	100.0
Total	Field-Dependent	33	41.2
	Field-Independent	47	58.8

Means, standard deviations, and t-test of respondents' GEFT scores by gender

GEFT scores of respondents were compared by gender (Table 3). It was found that the male students' GEFT mean score (mean = 13.18) was significantly higher than the female students' GEFT mean score (mean = 10.27).

As shown in table 3, the GEFT mean score of all respondents was 11.87 of a maximum possible score of 18.

Table 3. Means, standard deviations, and t-test of respondents' GEFT scores by gender

Variable	Total		Gender: Male		Gender: Female		t-Value
	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	
GEFT Scores	80	11.87 (4.25)	44	13.18 (4.17)	36	10.27 (4.30)	3.04**

Note: Raw scores are based on a maximum possible score of 18, ** p<.01

Correlation analysis between GEFT scores and age

Table 4 presents the results of the correlation analysis between GEFT scores and age of respondents. It is recognizable that age isn't correlated with the GEFT scores of respondents. Thus is can be concluded that there is no difference in learning style of students with different ages.

Table 4. Correlation analysis between GEFT scores and age

	GEFT Scores	Age
GEFT Scores	1	.203
Age		1

Means, standard deviations, and t-test of respondents' grade point average (GPA) by learning style

Means, standard deviations, and t-test of agricultural students' grade point average (GPA) by learning style are demonstrated in table 5. The data illustrates that the GPA for students preferred field-independent learning style (15.28 out of 20) was significantly higher than the GPA for students preferred field-dependent learning style (14.53 out of 20).

Table 5. Means, standard deviations, and t-test of respondents' grade point average (GPA) by learning style

Variable	Total		learning style				t-Value
			Field-Dependent		Field-Independent		
	n	Mean (SD)	n	Mean (SD)	n	Mean (SD)	
Grade Point Average (GPA)	80	14.97 (1.50)	33	14.53 (1.57)	47	15.28 (1.38)	-2.18*

Note: Range of GPA is between zero and twenty

* p<.05

Conclusions and Implications

One of the most significant challenges that university instructors face is to be tolerant and perceptive enough to recognize learning differences among their students. Many instructors do not realize that students vary in the way they process and understand information. While recognition of students' learning styles has massive benefits and implications for college admissions and for instructors, the main purpose of this study was to establish baseline information regarding the distribution of learning style among students, in order for colleges and individuals to benefit.

GEFT scores of respondents were compared by gender and showed that the male students' GEFT mean score was significantly higher than the female students' GEFT mean score. This result is accordant to the preliminary norm data on GEFT, in which college males performed significantly higher than college females (Witkin et al, 1971). Since not all students learn the same, it is imperative that instructors teach in a manner in which all learning styles are considered.

Since not all students learn the same; it is imperative that faculty members recognize the learning style differences of their students. They must pay attention to the fact that field-dependent students require externally defined goals, are extrinsically motivated, and need more explicit instruction in problem-solving. Conversely, field-independent students rely on self-defied goals, are more intrinsically motivated, tended to provide their own structure to facilitate learning, and prefer competition. Therefore, faculty members teach in a manner in which all learning styles are considered and incorporated.

According to results, learning style influences scholastic success of students. The GPA for students preferred field-independent learning style was significantly higher than the GPA for students preferred field-dependent learning style. So, learning style could be an important element in success in higher education. Hence, students must be assessed of their learning style. So, we recommend counseling on different learning styles and how agriculture students learn more effectively and efficiently, and how they adapt their learning style to various teaching styles in classrooms.

References

Cano, J. & Metzger, S. (1995). The relationship between learning styles and levels of cognition of instruction of horticulture teachers. *Journal of Agricultural Education*. 36(2): 36-42.

Davies, D. (1998). The virtual university: A learning university. *The Journal of Workplace Learning*, 10 (4), 175 – 213

Dembo, M.H. (1988). Applying educational psychology in the classroom (Third Edition). New York: Longman.

Garger, S. & Guild, P. (1984). Learning styles: The crucial differences, *Curriculum Review*, 23(1): 9-12.

Garton, B.L., Ball, A.L. & Dyer, J.E. (2002). The academic performance and retention of college of agriculture students. *Journal of Agricultural Education*, 43(1): 46-56.

Gregorc, A. F. (1979). Learning/teaching styles: Potent forces behind them. *Educational Leadership*. 36, 234-237.

Kolb, D.A. (1984). Experiential Learning Experience as the source of Learning and Development. Prentice-Hall, 1994. Carolina State University, USA.

Raven, M.R., Cano, J., Garton, B.L. & Shelhamer, V. (1993). A comparison of learning styles, teaching styles, and personality styles of pre-service Montana and Ohio agriculture teachers. *Journal of Agricultural Education*. 31(1): 40-50

Shih, C. & Gamon, J.A. (2001). Relationship among learning strategies, patterns, styles, and achievement in Web-based courses, *Journal of Agricultural Education*. 42(4): 12-20.

Weisburg, M. & Ullmer, E. J. (1995). Distance learning revisited: Life-long learning and the National Information Infrastructure. *Proceedings of the 1995 Annual National Convention of the Association for Educational Communications and Technology* (AECT), p. 628-647.

Witkin, H.A. (1973). The role of cognitive style in academic performance and in teacher-student relations. *Research Bulletin*, Educational Testing Service, Princeton, NJ, 73-101.

Witkin, H.A., Oltman, P.K., Raskin, E., & Karp, S.A. (1971). *Group Embedded Figures Test Manual*. Palo Alto, CA: Consulting Psychologist Press.

STUDENTS' DEVELOPMENT OF CONCEPTUAL KNOWLEDGE WITHIN THE TOPICS THERMAL EQUILIBRIUM AND HEAT TRANSFER

Christian Rogge
Justus Liebig University Giessen

Abstract

In conceptual change research, a common procedure for the investigation of learners' conceptions is to search for "similar" (intra- or inter-individual) utterances and activities in order to infer an underlying conceptual structure. However, there is still a lack of precise criteria that describe what actually constitutes a conception and how conceptions are assigned to learners' activities and thinking processes. In order to address this problem, this study operationalises a conception as follows: learners' verbal activities are regarded as conceptual if they comprise generalised considerations which do not refer to a particular situation. In addition, conceptual knowledge is only assigned if the generalisation is expressed explicitly by the learner. Based on this approach a coding schema was developed with which the (conceptual) quality of learners' situated knowledge can be assessed. Video-data from students of grade 8 and 11 working in groups on instruction about the topic heat transfer were used to describe the development of students' knowledge and their conceptions during instruction. Results show both a basically low frequency of explicit conceptions and a large time requirement for developing a conception.

Rationale

In conceptual change research, learners' conceptions are typically analyzed by using questionnaires or interviews (e.g., Duit, 2008). Related research studies usually focus on identifying a limited number of different (intermediate) conceptions (e.g., Slotta, Chi & Joram, 1995; Vosniadou & Brewer, 1992). A common procedure is to search for "similar" (intra- or inter-individual) utterances and activities in order to infer an underlying conceptual structure. However, there is still a lack of precise criteria that describe what actually constitutes a conception and how conceptions are assigned to learners' activities and thinking processes (e.g., diSessa & Sherin, 1998). In order to address this problem, this study operationalises a conception as follows: learners' verbal activities are regarded as conceptual if they comprise generalised considerations which do not refer to a particular situation. This perspective on conceptual knowledge is supported by arguments claiming that conceptual knowledge refers to an "implicit or explicit understanding of the principles that govern a domain […]." (Rittle-Johnson, Siegler & Alibali, 2001). In contrast to the common procedure in the literature, in this study, a conception is only assigned if the generalisation is expressed explicitly by the learner. Another difference is that the ascription of conceptions is carried out on the basis of rather short time scales (not longer than 10 seconds). Based on this process-based approach, the aim of the study is to shed new light on conceptual change research. Therefore, the project explores the following research questions:

1. Which criteria are useful to describe students' development of conceptual knowledge?
2. How do students develop a conceptual level of understanding?
3. How often do students make use of conceptual knowledge (of different qualities) while working on physics tasks?

Methods

Sample and Procedure

The study reported here is part of a wider research project that investigates how students' argumentation interrelates with their scientific knowledge (for details see v. Aufschnaiter et al., 2008). During the main study 18 students from grade 8 (about 13 years old) and 12 students from grade 11 (about 16 years old) worked in groups of three on tasks about the topics electric circuits, heat transfer, light & shadow and blood pressure. Every group worked on two out of four topics, so every topic was addressed by three groups from grade 8 and two groups from grade 11. All student activities were recorded on video. For every topic a learning unit was developed that comprises two sessions each lasting about 60 to 80 minutes with a break of one week in-between. All units were designed similar to a teaching experiment, in which students can work in small groups independently on the material. All instructions were written on cards and material for carrying out experiments was provided. With this design, all interactions with the groups were controlled so that processes of the groups could be compared. The design of the concrete learning activities was informed by the Model of Educational Reconstruction (e.g., Duit, Gropengießer, & Kattmann, 2005) and by results on students' learning processes in physics (e.g., v. Aufschnaiter, 2006). The starting point and main focus of the instruction is on providing students with systematic experiences about specific phenomena and situations. For each part of a topic, sequences of tasks were designed consisting mainly of experimental activities in order to help students to discover by themselves the principles of the presented phenomena (see also v. Aufschnaiter, 2006). At the end of a sequence students were asked to find more general descriptions and explanations for the phenomena they had observed. Information about the principles of the phenomena and their (theoretical) explanation were also offered at the end of a sequence of tasks. During further sequences students got the possibility to "rediscover" conceptions several times.

Data Analysis

For this project, video-data on one unit (heat transfer) are taken into account, corresponding to about 14 hours of videos (video-data on the other learning units are analyzed by other members of the research group). In order to investigate the relatively expanded amount of video-data, analyses were performed in two successive steps: Firstly, category-based analyses of videos and, secondly, in-depth investigations of transcripts.

Category-based Analyses of Videos

In order to get an overview about the data and to generate quantitative results all data concerning the topic heat transfer were coded directly at the video. For that purpose, category systems were developed with respect to low inference variables (students' general activities, area of content) and with respect to high inference variables (e.g., level of students' conceptual understanding). All codings were carried out on the basis of time sampling (analysis unit: 10 seconds) using the software Videograph. Every student of a group was coded individually. The category system for the description of students' conceptual understanding was developed on the basis of an existing coding schema (v. Aufschnaiter, 2006). This original coding schema was tested with selected parts of the data in order to examine whether the categories could be successfully applied to the activities of the participating students. As a result, some categories were (partly) refined. The resulting category system is presented in Table 1. In its main idea, the system distinguishes between explorative, intuitive rule-based and explicit rule-based activities. Only verbal activities of the last kind referring explicitly to the commonalities of several objects or situations are considered as being conceptual. Each area is further differentiated into three subcategories (see Table 1). Controls of intercoder reliability yielded acceptable values of Cohen's Kappa from 0.55 to 0.68, taking into account the high inference of the coding system.

Table 1. Categories for the description of students' (conceptual) understanding

Main categories	Subcategories	Description	Example (topic: heat transfer)
1: explorative approach	act/ experiment	Students explore phenomena, e.g., carry out an experiment or measure a value. In addition, students can simultaneously describe their activity. [Just watching, reading or writing is not coded.]	(student touches an iron cube) "Touch this iron cube. It's cold."
	describe with visual aid	Students observe objects, activities or situations and describe them.	(student looks at the thermometer) "The temperature is increasing."
	describe without visual aid	Students describe objects, activities or situations without observing them. Also: Students make a guess what will happen.	[student remembers:] "The water got colder."
2: intuitive rule-based approach	assume	Students make an assumption about what will happen. Students emphasize an aspect that is important from their point of view.	"The cold water in the Petri dish will certainly reach 22 degree."
	attribute	Students make use of specific linguistic elements (particularly physics terms) to label and describe phenomena and objects.	"This hot gel pack is a heat source."
	explain	Students explain how different concrete aspects, phenomena or situations relate to each other.	"This gel pack didn't cool down because it's wrapped in a newspaper."
3: explicit rule-based approach (conceptual)	generalise	Students express a generalisation explicitly. Students formulate a rule-based relationship.	"Objects adapt to the temperature of the environment."
	explain rule-based	Students use generalisations or rule-based relationships in order to explain a particular or general situation.	"The rod has room temperature because objects adapt to the temperature of the environment."
	predict rule-based	Students explicitly refer to generalizations or rule-based relationships when predicting the progress of a particular or general situation (e.g., the result of an experiment).	"The white sheet of paper won't get that warm because light and bright surfaces reflect thermal radiation."

In-depth Investigations of Transcripts

The procedure for transcript-based analyses started with the identification of all student activities referring to a specific content area, e.g., all activities and statements that are related to thermal equilibrium. Then, students' development of conceptual knowledge was traced in the course of time, separately for each specific content area. Students' conceptual development was also investigated by means of additional criteria: students' thematic focus, the number of different content elements that are integrated into students' considerations and whether or not students' conceptions are inferred directly from experiences (phenomenon-based) or are based on theoretical assumptions (model-based).

Results

Category-based Analyses of Videos

On average, each student showed an active engagement with the contents of the tasks (e.g., carrying out experiments and/or talking about physics content) in about 31% of total time (see Figure 1). During the rest of the time each student performed other activities like observing/listening (35%), organisational activities (e.g., gathering or putting away materials) (21%) and reading/writing (7%). The individual proportion of active participation varies substantially between the students (15-43%). No significant differences between grade 8 and 11 were found. The categories for the description of students' conceptual understanding were only applied to situations where a student showed an active participation. On average, each student rarely expressed explicit conceptualisations (5% of coded events, see Figure 1). Instead, students basically followed an explorative (54%) and intuitive rule-based approach (41%). Here, too, large differences exist between the students (e.g., explicit conceptualisations vary from 1-10%). While students' active participation does not differ significantly between grade 8 and 11 (see above), differences with respect to the quality of these activities were found. The students from grade 11 showed a significant higher proportion (46%) of intuitive rule-based activities than the students from grade 8 (36%). In particular, the 11th graders explained more often (10%) how different concrete aspects relate to each other than the 8th graders (5%). Finally, it is noticeable that the proportion of explicit conceptualisations is roughly the same in grades 8 and 11. Since findings from other studies also show that students seldom reach a (deeper) conceptual understanding (e.g., TIMSS, PISA), it is concluded that the mentioned results are typical for students and not an artefact of the instruction used.

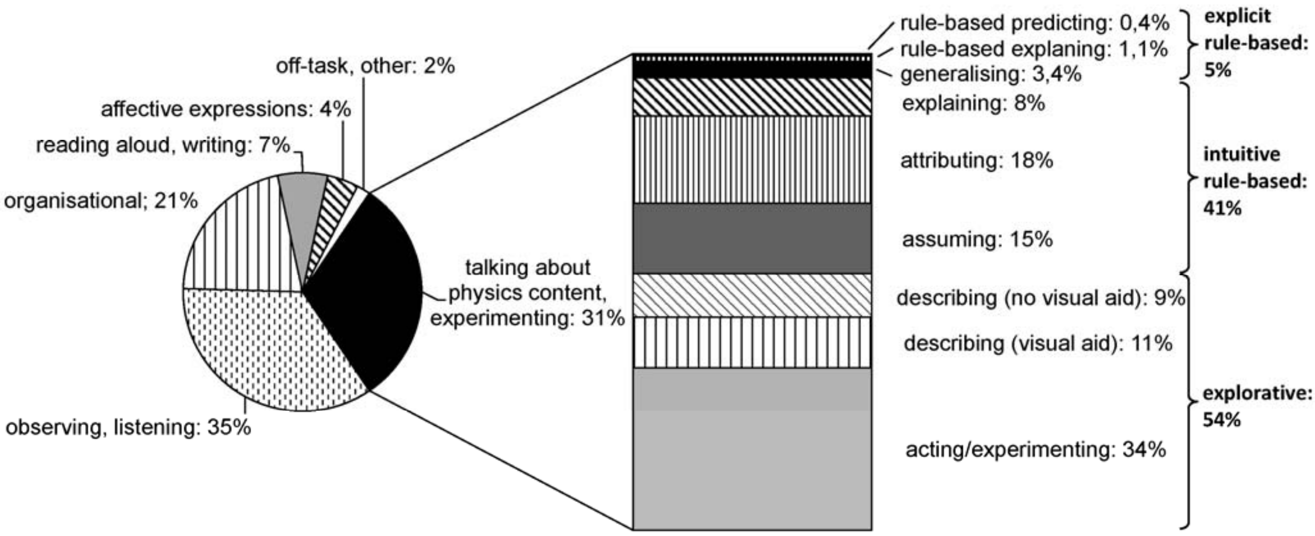


Figure 1. Frequency of activities (100% = total duration of the unit heat transfer ≈ 170min, mean values for 12 students) and distribution of conceptual categories (100% = total duration of content-specific activities)

In-depth Investigations of Transcripts

For each student his/her development of (conceptual) knowledge within a specific content area was traced in the course of time. The analyses show that students' situated understanding developed from an explorative to an intuitive rule-based approach, especially during the beginning of a newly presented content area. Explicit conceptualisations typically occurred only after various and systematic experiences with the content area and if prompts asked for more general aspects of the content. This general development from an explorative approach to an explicit rule-based (conceptual) approach is not a simple linear progression, but occurs in a more cyclical manner, starting with repeated movements from an explorative to an intuitive rule-based level of understanding (see Table 2 for an example).

Table 2. Example of a conceptual development for a group from grade 8 working on a sequence of tasks concerning the adaption of objects' temperatures to the temperature of the environment

Time	Context and students' activities (shortened, grade 8)	Conceptual level
0:00:00	**Task 1.1-1.3:** S(tudents) sort objects, whether they feel warm, normal, or cold.	1 explorative
	→ While searching for another „warm" object: *"No, not this one, it is metal."*	2 intuitive rule-based
0:15:10	**Task 1.4-1.6:** S measure & compare the objects' temperatures (all at ~24°C)	1 explorative
	→ *"This can't be!" – "The iron cube can't be so warm like the wooden one!"*	2 intuitive rule-based
0:33:00	**Task 1.7:** Measure the room temperature, compare it with the measured ones.	
	→ S measure 23,5°C.	1 explorative
	→ *"It* [the objects' temperature] *adapts to the room."*	2 intuitive rule-based
0:37:00	**Task 1.8:** Temperatures of four cubes of different materials laying in a fridge?	
	→ S reckon: cube made of aluminium, iron, stone: 5°C; wooden cube: 10°C	2 intuitive rule-based
0:38:20	**Task 1.9:** Measure the air temperature in the fridge & the cubes' temperatures	
	→ S measure the temperatures (all at ~4-5°C) – astonishment	1 explorative
	→ *"They* [the objects] *adapt always to the environment."*	3 explicit rule-based
0:41:30	**Task 1.10/1.11 :** Progression of the temperatures of hot and cold water?	
	→ S measure the temperatures and observe warming and cooling	1 explorative
	→ *"I know why this happens." – "Yes, because the water adapts to the air."*	2 intuitive rule-based
0:49:20	**Task 1.12:** Progression of the temperature of a knife brought out of a fridge?	
	→ S measure the temperature of the knife: *"It increases more and more."*	1 explorative
	→ *"Sure, it will adapt to the air."*	2 intuitive rule-based
0:52:00	**Task 1.13:** What is the temperature the hot water, the cold water, and the cold knife will finally get to? → *"It adapts always to the air temperature."*	3 explicit rule-based

Conclusions and Implications

The presented methodological approach provides an alternative to assess the (conceptual) quality of learners' situated activity. Results show both a basically low frequency of explicit conceptualisations and a large time requirement for developing an explicit conceptualisation. This indicates that conceptual change research seems to overestimate the role of (mis)conceptions in "producing" activity and to underestimate the role of experiences in establishing a conceptual understanding.

References

diSessa, A. A., & Sherin, B. L. (1998). What changes in conceptual change? *International Journal of Science Education, 20*(10), 1155-1191.

Duit, R. (2008). *Bibliography - STCSE (Students' and Teachers' Conceptions and Science Education).* Kiel: IPN. Online: http://www.ipn.uni-kiel.de/aktuell/stcse/stcse.htm.

Duit, R., Gropengießer, H., & Kattmann, U. (2005). Towards science education research that is relevant for improving practice: The model of educational reconstruction. In H. E. Fischer (Ed.), *Developing standards in research on science education* (pp. 1-9). Leiden: Taylor & Francis.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology, 93*(2), 346-362.

Slotta, J. D., Chi, M. T. H., & Joram, E. (1995). Assessing students' misclassifications of physics concepts: An ontological basis for conceptual change. *Cognition and Instruction, 13*(3), 373-400.

von Aufschnaiter, C. (2006). Exploring the processes of students' development of physics concepts. *Paper presented at the conference of NARST*, San Francisco, USA, April 2006. (CD-Rom)

von Aufschnaiter, C., Fleischhauer, J., Rogge, C., & Riemeier, T. (2008). Argumentation and scientific reasoning - An exploration of their interrelationship, *Paper presented at the conference of NARST*, Baltimore, USA, April 2008. (CD-Rom)

Vosniadou, S., & Brewer, W. F. (1992). Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology, 24*(4), 535-585.

WHAT BROUGHT THEM INTO SCIENCE? UNIVERSITY STUDENTS' INTEREST PROFILES AND MOTIVATION FOR CHOOSING SCIENCE AT UNIVERSITY LEVEL

Berit Bungum
Norwegian University of Science and Technology

Silje Rødseth
Norwegian University of Science and Technology

Annette Lykknes
Norwegian University of Science and Technology

Kari Mette Murvoll
Norwegian University of Science and Technology

Abstract

A survey is undertaken in order to elicit the interest profiles and motivation for going into science among university students of biology, chemistry and physics. Comparisons are made to other studies among younger students in secondary school. In a questionnaire, students were asked to rank topics from school physics, chemistry and biology respectively; with regards to how interesting they had experienced them. Results indicate that the students' main motivation is to develop their own knowledge in a field intriguing to them. Some differences between students of the three subjects were found. Philosophical aspects of the subject are of high interest to physics students, while chemistry students prefer laboratory and industry related topics. Biology students show the highest variation in interest profiles, and this is also the only group where gender differences can be detected. Female biology students seem to be attracted to modern biology on micro level while males appear to prefer the field related parts of the subject. A majority of students of all subjects report research as their main career aspiration. This fits well with their focus on interest in choosing their lines of study, but may also indicate that the wide range of career possibilities within the natural sciences should be better communicated in school science.

Introduction

Recruitment to the natural sciences in higher education is a matter of extensive concern. Various studies have indicated that studying science is not consistent with how Western youth wish to construct their identity (Hannover & Kessels, 2004; Schreiner & Sjøberg, 2007). While science is often seen as a collection of static and true facts to be learnt, young people of today prefer careers that to a higher degree foster individuality, self development and creativity.

Although many young people do not see further studies and careers in science as a relevant option, they do show interest in science. Several studies have investigated what science topics pupils in secondary education find interesting. The ROSE study has shown gendered patterns among 15 year-olds; while boys find technology and spectacular phenomena interesting, girls' interests are in the direction of health issues and "mysterious" phenomena (Schreiner & Sjøberg, 2007). The Norwegian FUN study, investigating the interests and attitudes of physics students in upper secondary school, showed that students find topics that relate to philosophical issues, such as astrophysics and relativity most interesting (Angell, Guttersrud, Henriksen & Isnes, 2004). In contrast to the ROSE study, only small gender differences were found in the interest profiles of these older students who chose to study physics in school. Other studies indicate, not surprisingly, that students opt for the sciences in school for a variety of reasons, not only because they want to become scientists (Cleaves, 2005).

Little research is, however, undertaken to get insight into the interests and motivations among those students who actually choose to start on university studies in the natural sciences, and hence going towards a science-related career. The present study investigates what these students found interesting in the subjects in school, their motivation for choosing to study physics, chemistry and biology respectively, and how they look upon their future careers in science.

Research questions

The research questions for the study were the following:

- What are the interest profiles of the science students?
- What do science students see as important for their future career?
- Are there differences between fields of study or between genders in this regard?

Methods

The study was undertaken as a questionnaire survey among young students in physics, chemistry and biology courses at The Norwegian University of Science and Technology. The questionnaire listed topics adapted to the Norwegian curriculum of physics, chemistry and biology for upper secondary school. Students were asked to state on a 4 point Likert scale to what degree they found various elements of the subject interesting in school and to what degree various reasons for their choice were important. They were also asked how important they consider characteristics of a job to be for their career choice. These questions were identical to a selection of questions appearing in the ROSE questionnaire to 15 years olds (see Schreiner & Sjøberg, 2007) in order to facilitate comparisons between the age groups. The job characteristics included: Working with something I find important and meaningful, having plenty of time for my family, having a good income, becoming a leader at work, and developing my knowledge and abilities.

The questionnaire was administrated during lectures in the respective subjects, ensuring a high response rate from the students. 288 students answered the questionnaire, whereof 142 were physics students, 84 chemistry students and 62 biology students, distributed among study programmes of engineering (chemistry and physics), general science (biology, chemistry, physics) and biotechnology. Students enrolled in the latter programme were asked to choose whether their main identity was as chemistry students, or within biology. The distribution on subjects and gender in the sample is given in Table 1.

Table 1. Distribution on subjects and gender in the sample.

	Physics	Chemistry	Biology	Total
Females	59	51	49	159
Males	83	33	13	129
Total	142	84	62	288

Results

What are the science students' interest profiles?

A majority of students of all groups motivate their choice of study and career due to interest in the subject. Figures 1-3 present results on what aspects of the subjects in school the students found interesting, and that hence have inspired them in their choice of studying the subject at university. The figures display results for women and men separately.

As shown I Figure 1, the topics found most interesting by physics students are astrophysics, relativity and quantum mechanics. This result resembles what has earlier been found in upper secondary school (Angell et al., 2004), that students are motivated by physics topics that bring the subject in the direction of philosophical issues rather than in a utilitarian direction. Experiments in physics and topics with links to applications and technology get lower scores. Remarkably, this profile of interest applies to the students enrolled in the technologically oriented engineering program as well as to those in the more general and scientific oriented science program. Nor did we find any important gender differences in the interest profile among the physics students.

Chemistry students show a different profile than their peers in physics. These results are shown in Figure 2. Doing experiments, and also laboratory and industry oriented topics, such as organic chemistry and chemical analysis, get the highest scores. The more theoretically oriented (and demanding) topics, such as electro-chemistry and acids/bases are less popular among the chemistry students. This profile applies to both males and females, and to students in the different programs of study.

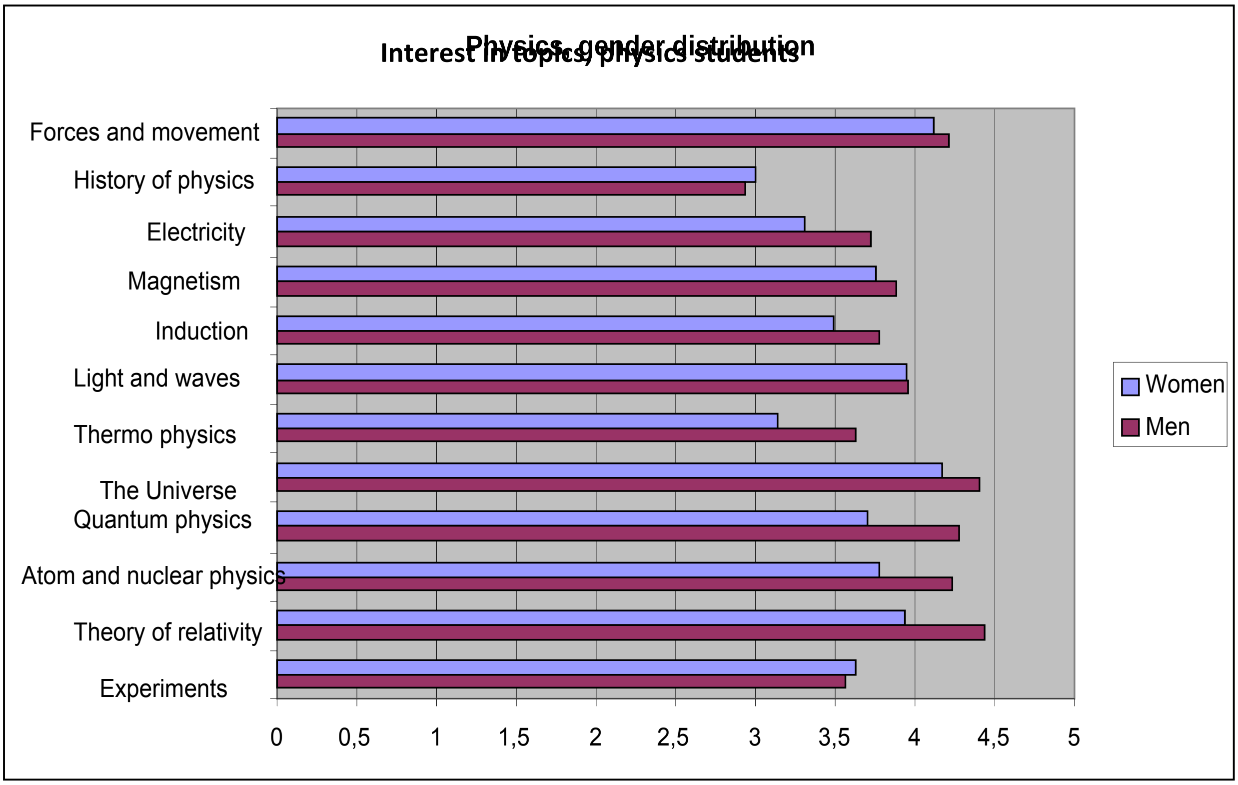


Figure 1. Means for physics students' interest in physics topics. The scale runs from 0 = "Not interesting" to 5 = "Very interesting".

Results showing the interest profile of biology students are shown in Figure 3. Seen as a group, they report that their main interest is within modern biology such as cell biology, genetics and gene technology. This might be attributed to the fact that one quarter of these students are enrolled in a program for biotechnology studies. However, the cell biology scores high among all biology students. Among students in the general science program, topics related to environment also score high together with doing experiments.

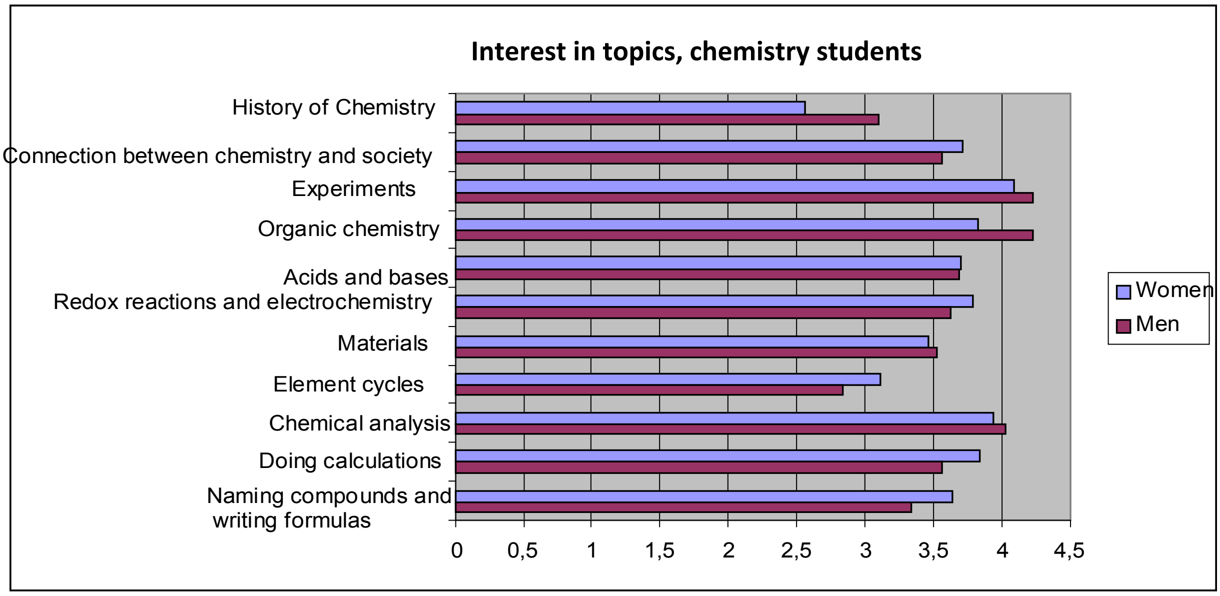


Figure 2. Means for chemistry students' interest I chemistry topics. The scale runs from 0 = "Not interesting" to 5 = "Very interesting".

Furthermore, results from the biologists show the strongest gender pattern in our survey. The biology programs are dominated by female students, and these hence dominate the means referred above. Looking at the two genders separately reveals a different interest profile for males. While female biology students express most interest in cell biology and genetics, their male peers express that they find doing experiments, field investigations and ecology most interesting. This might indicate that female students in biology are oriented towards the modern parts of biology as a science, while the male biology students go in more traditional directions within the subject.

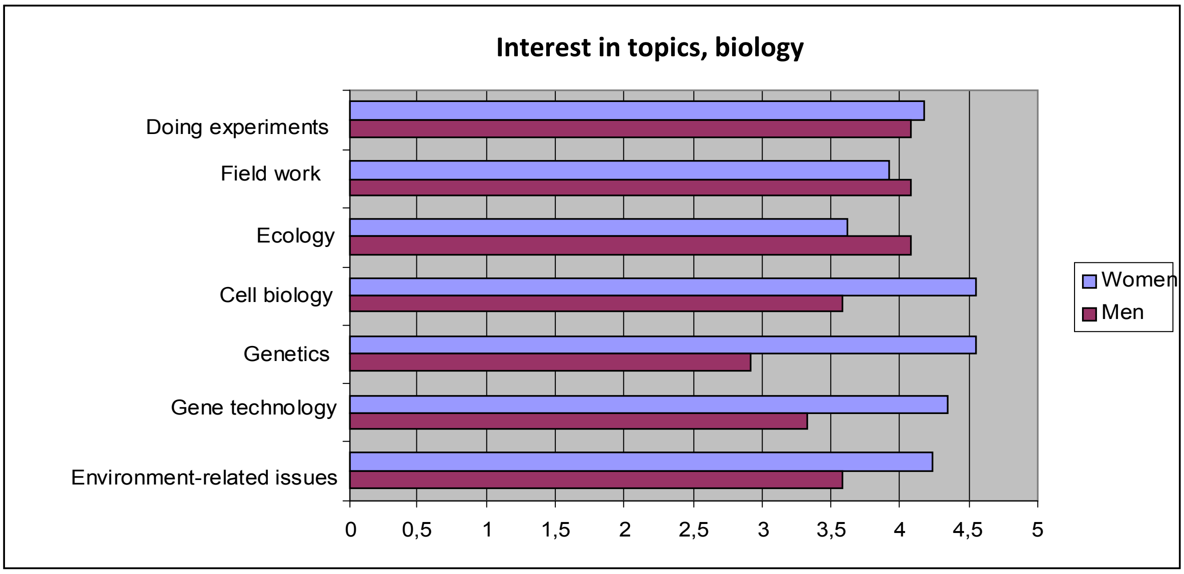


Figure 3. Means for biology students' interest in biology topics. The scale runs from 0 = "Not interesting" to 5 = "Very interesting".

The questionnaire also contained questions about how interesting they find core characteristics of their subjects. Science as a field of study may appeal to students due to its philosophical and curiosity-driven aspects, or by its technical or utility-driven aspects. The former dimension was operationalized as "How nature works" for all students, while the latter was presented in different ways. For Physics students it was described as "technical physics" and for Chemistry students "Chemical analysis, industry and production". Biology students were presented to two categories in this regard; "Biotechnology" and "Cell biology/physiology". Results are given in Figures 4-6. They show that as groups students in all the three disciplines express more interest in the curiosity-driven aspect of scienc than in technical aspects or applications. The only gender difference found in this regard is between girls and boys within biology; girls seem to have their interest more towards biotechnology than what do boys.

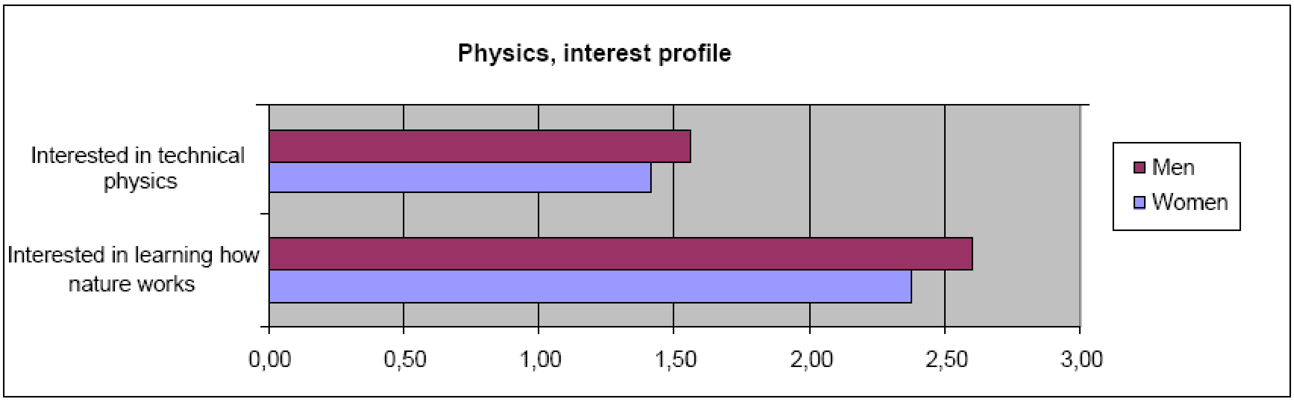


Figure 4. Means for physics students' responses to interest in learning about technical physics and how nature works. The scale runs from 0 = "Do not agree" to 4 = "Completely agree".

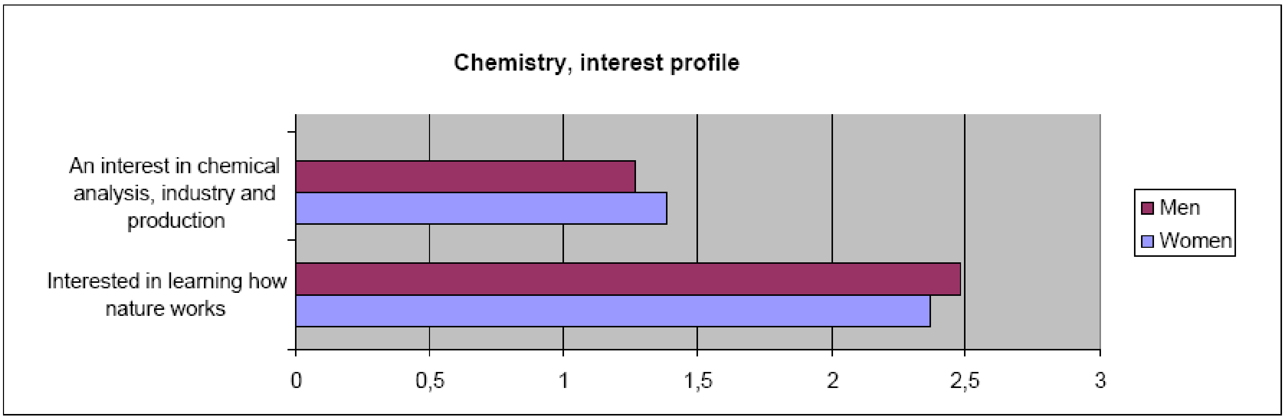


Figure 5. Means for chemistry students' responses to interest in chemical analysis, industry and production, and how nature works. The scale runs from 0 = "Do not agree" to 4 = "Completely agree".

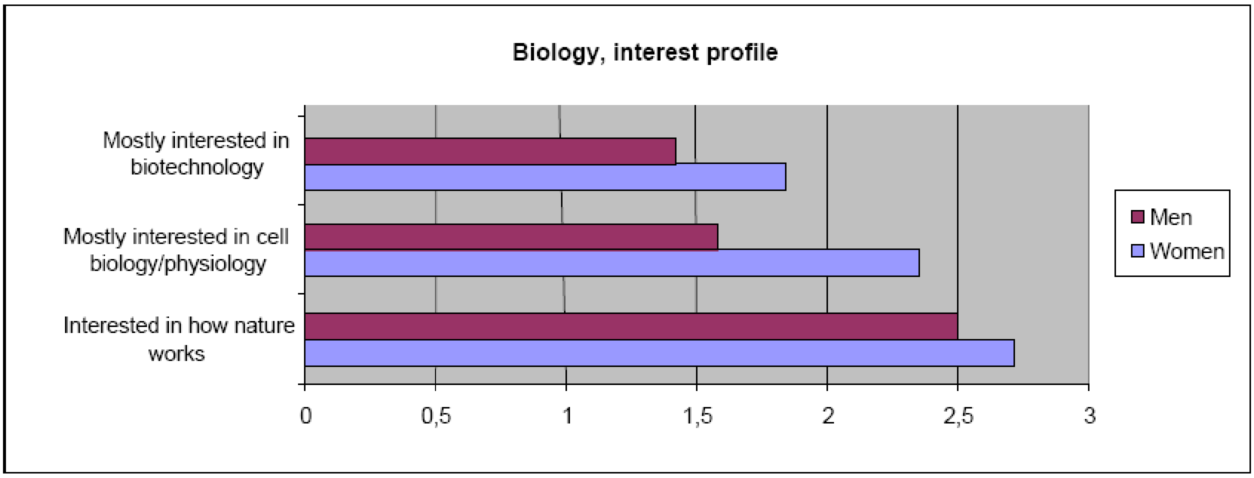


Figure 6. Means for biology students' responses to interest in biotechnology, cell biology/physiology, and how nature works. The scale runs from 0 = "Do not agree" to 4 = "Completely agree".

What do science students see as important for their future career?

The students' responses to questions about their wishes for a future job resemble to some degree those from 15 year olds from the ROSE study. As shown in Figure 7, the students, regardless of gender and field of study, see it as essential to work with something they find important and meaningful. They also value very high to develop their own knowledge and abilities. Becoming a leader and having a good income, however, is less valued by the university science students than by their younger peers. While the ROSE study showed a noteworthy gender difference on pupils' interest in "Working with people rather than things", the responses in our study indicate that this factor is not gendered among university science students. Nor does it seem to vary according to field of study, quite contrary to the common anticipation that biology students are more human-oriented than their peers in physics.

When asked what kind of work they wished to have in the future, most students of all three subjects and of both genders point to research as their first choice. This is followed by work in private companies. Going into teaching, health-related work or societal administration and management are seen as far less appealing options. These results are shown in Figure 8.

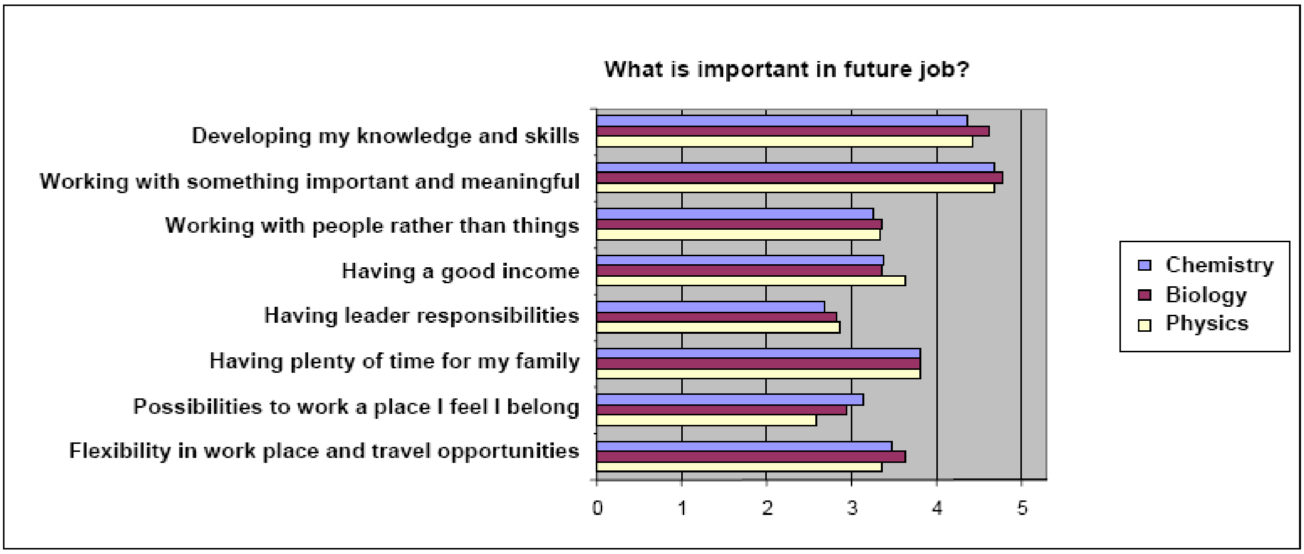


Figure 7. Students' responses to what they see as important in their future job. The scale runs from 0 = "Not interesting" to 5 = "Very interesting".

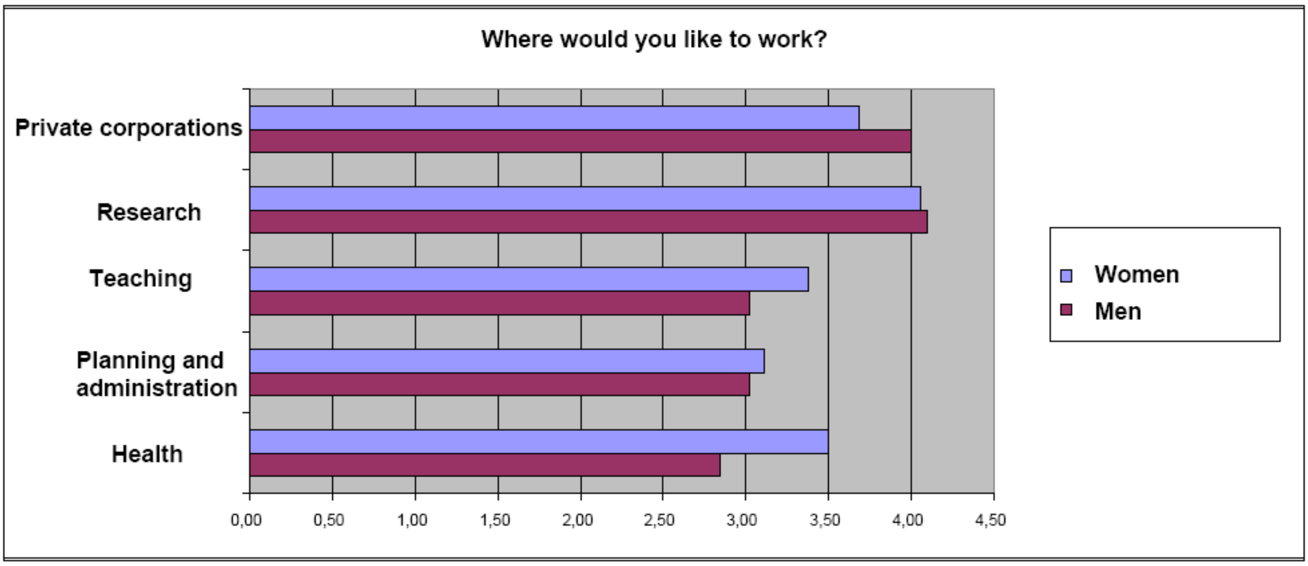


Figure 8 . Means for students' responses to where they would like to work. The scale runs from 0 = "Not interesting" to 5 = "Very interesting".

Conclusions and Implications

Results from the study suggest that university students in the natural sciences do not differ from modern youth in general in that their choice of study comprises a part of their construction of identity. Developing their own knowledge and working with something meaningful are regarded as highly important for their future work. Gender differences in responses about interests and career choice are far less prominent than among pupils in secondary school. This might either be due to maturation, or to the fact that the university students in the sciences represent a more narrow population than in school science. Further research is required in order to shed more light on this issue.

Interest profiles show considerable differences between students of various subjects. While physics students, even those in a technology program, respond a high interest in philosophical aspects of the subject, chemistry students are more content with laboratory-related aspects.

Bology students are those showing most variation, with regards to gender and study program, in what elements of the subject they find most interesting. It seems, however, that it is mainly modern biology and applications that has attracted the majority of these students to biology.

All the students see research as the most appealing field of working in their future career as scientists. This resembles their focus on interest and knowledge development in choosing their field of study. With regards to recruitment, however, this might mean that there is a need in secondary education for a wider exposure of the variety of career possibilities students have with university education in one of the natural sciences.

References

Angell, C., Guttersrud, Ø., Henriksen, E. K. & Isnes, A. (2004). Physics: Frightful, But Fun. Pupils' and Teachers' Views of Physics and Physics Teaching. *Science Education, 88*(5), 683-706.

Cleaves, A. (2005). The formation of science choices in secondary school. *International Journal of Science Education, 27*(4), 471-486.

Hannover, B., & Kessel, U. (2004). Self-to-prototype matching as a strategy for making academic choices. Why high-school students do no like math and science. *Learning and Instruction, 14*(1), 51-68.

Schreiner, C. & Sjøberg, S. (2007). Science Education and Youth's Identity Construction – Two incompatible projects? In D. Corrigan, J. Dillon R. Gunstone (Eds.), The Re-emergence of Values in the Science Curriculum (pp 231-247). Rotterdam: Sense Publishers.

DIFFERENCES IN 10TH GRADE STUDENTS' USE OF SELF-REGULATORY LEARNING STRATEGIES BY SCHOOL TYPE

Cansel Kadioglu
Gaziosmanpasa University

Esen Uzuntiryaki
Middle East Technical University

Abstract

This study aimed to investigate whether students attending different high schools differed in their use of self-regulatory learning strategies in chemistry course. The learning strategies studied in this study were rehearsal, elaboration, organization, critical thinking, metacognitive self-regulation, time and study environment, effort regulation, peer learning, and help seeking. Three hundred fifty two 10th grade students from different high schools (Anatolian High School vs. Regular High School) participated in the study. Learning strategies scale of Motivated Strategies for Learning Questionnaire (MSLQ) was administered to the students. Multivariate Analysis of Variance (MANOVA) was used, assigning nine strategies as dependent variable and school type as independent variable. Results revealed statistically significant difference between school types on combined dependent variables. Univariate comparisons revealed significant differences between two school types on five strategy types (rehearsal, elaboration, organization, critical thinking, and metacognitive self regulation). Students attending regular high schools were reported to use these strategies more often.

Introduction

Self-regulated learning (SRL) has become an important topic in education for nearly three decades. Researchers propose different theories (such as social cognitive, Vygotskian, and cognitive constructivist approaches) to explain features of SRL (Zimmerman, 2001). These theories emphasize some common features: Students are aware of the processes that improve their academic achievement and they monitor these processes by getting feedback from earlier learning processes; these processes follow a cyclic way; and SRL includes motivational processes. Social cognitive theory is the most commonly studied one among these theories. Social cognitive theory explains human functioning through reciprocal interactions between personal (student's self-efficacy beliefs), environmental (feedback from the teacher), and behavioral (attention towards the instruction) factors (Bandura, 1986). In addition, it emphasizes the *agency* of the learner which refers to the control of individuals over their learning processes. According to this view, the learner makes his/her own choices and continues his/her learning with respect to these choices in order to achieve his/her own goals. Based on social cognitive theory, Zimmerman (2000) defines SRL as "self-generated thoughts, feelings, and actions that are planned and cyclically adopted to the attainment of personal goals" (p.14).

Effective use of self-regulatory strategies activates students' learning in many ways. Self-regulated learners determine their learning goals, select the most appropriate strategies, are responsible for applications of these strategies, accept teacher guidance when necessary, and are aware of their strengths and weaknesses. In addition, they monitor their learning process, observe their progress, and evaluate whether they achieve their goals or not (Zimmerman, 1990). Recent studies indicate a positive association between students' SRL skills and their academic achievement (Pintrich & DeGroot, 1990; Pape & Wang, 2003; Yumusak, Sungur, & Cakiroglu, 2007).

Indeed, self-regulated learners not only improve their academic achievement, they also become aware of what they know and what they do not know. Highly self-regulated learners also know the strategies that improve their learning and use these strategies to achieve their goals. Zimmerman and Martinez-Pons (1986, p.615) define SRL strategies as "actions directed at acquiring information or skill that involve agency, purpose (goals), and instrumentality self-perceptions by a learner". In the literature investigating students' self-regulatory skills, Motivated Strategies for Learning Questionnaire (MSLQ) developed by Pintrich, Smith, Garcia, & McKeachie (1991) is commonly used around the world. MSLQ assesses students' motivational orientation and learning strategies associated with SRL. The focus of the present paper is students' learning strategies. Researchers defined nine categories of learning strategies and grouped these strategies for the most part in two categories: Cognitive and metacognitive strategies as one category and management of different resources as the other. Cognitive and metacognitive strategies include rehearsal, elaboration, organization, critical thinking and metacognitive self-regulation strategies. *Rehearsal strategy*, which refers to reciting or naming parts to be learned from the material, is generally used for simple tasks. It is associated with working memory and used during the encoding processes. *Elaboration strategy* is associated with long term memory and used to make connections between new information and prior knowledge. Paraphrasing, summarizing or creating analogies can be given as examples for this category. *Organization strategy* can be defined as making connections among the material to be learned. Outlining and selecting the main ideas in a text are examples for this category. *Critical thinking* occurs when students apply their previous knowledge in new situations in order to solve problems, make decisions or evaluations. Lastly, *metacognitive self-regulation strategies* include strategies such as planning, monitoring and regulation to increase awareness or gain control over one's cognition. On the other hand, resource management strategies are labeled as time and study environment, effort management, peer learning and help seeking. Management of *time and study environment* includes strategies for using time effectively and organizing the study environment to avoid distracters. *Effort management* refers to ones' management of effort to accomplish a goal in the face of difficulties and distracters. *Peer learning* means collaboration with peers during the learning process. Help seeking refers to asking for assistance when help is needed. Highly self-regulated learners are believed to use these strategies frequently.

Based upon aforementioned literature this study aimed to investigate whether students attending different types of high schools differed in their use of self-regulatory learning strategies in chemistry class.

Methods

Subjects of the Study

A total of 352 tenth grade students enrolled in chemistry courses at public high schools in Turkey participated in the study. The data were gathered from 122 tenth grade students from an Anatolian high school, and 230 students from two regular high schools. In Turkey, students take a Level Determination Examination (SBS) at the end of each grade level (6^{th}, 7^{th} and 8^{th}) at middle school since 2007-2008 academic year. Each student is placed in a high school according to his or her total score, in consideration of all three examinations taken after each grade. However the participants of the present study carried out a different process; they had to pass one examination at the end of 8^{th} grade called Secondary Schools Examination (OKS) in order to attend Anatolian high schools. These students did not need to take this examination to attend regular high schools. Therefore, students in Anatolian high schools are expected to have higher achievement than students in regular high schools. The chemistry curriculum is the same in both types of high schools. In addition, higher importance is given to second language education at Anatolian high schools.

Instrument

Learning strategy section of Motivated Strategies for Learning Questionnaire (MSLQ), was used to assess students' use of different cognitive and metacognitive strategies (rehearsal, elaboration, organization, critical thinking, metacognitive self regulation), and management of different sources (time and study environment, effort regulation, peer learning, and help seeking). The instrument was originally developed by Pintrich et al. (1991), and translated and adapted into Turkish by Sungur (2004). It is a self-report questionnaire where students rate themselves on a seven point Likert scale from "not at all true for me" to "very true for me".

To assess the factor validity of the scale, Confirmatory Factor Analysis using LISREL 8.30 for Windows (Jöreskog & Sörbom, 1993) was performed considering the factor structures proposed in the original scale. When the fit statistics for a nine factor solution were checked, it was found that the chi-squared to degrees of freedom ratio (χ^2/df) was 3.86; the goodness of fit index (GFI) was 0.73; and the root mean residual (RMR) was 0.08. These values were close to the original version of the scale and they were satisfactory and reasonable (Pintrich et al., 1991).

The reliability analyses were conducted for the nine subscales. The Cronbach's Alpha coefficients of the subscales were ranging from 0.48 for "effort regulation" to 0.82 for "critical thinking" and "metacognitive self-regulation". Table 1 presents the Cronbach's alpha values for the questionnaire.

Table 1. Reliability Coefficients and Number of Items for Subscales.

Subscale	Number of Items	Reliability coefficients
Rehearsal	4	.77
Elaboration	6	.81
Organization	4	.76
Critical thinking	5	.82
Metacognitive self-regulation	12	.82
Time and study environment	8	.72
Effort regulation	4	.48
Peer learning	3	.61
Help seeking	4	.59

Analysis of Data

Multivariate Analysis of Variance (MANOVA) was used to investigate whether students attending Anatolian versus regular high schools differed in their use of self-regulatory learning strategies. Nine strategies were assigned as dependent variable and school type was assigned as independent variable.

Results

Descriptive statistics (the mean scores and standard deviations) of each dependent variable across two school settings are given in Table 2. It was found that students attending regular high schools perceived that they used metacognitive self regulation, elaboration, and rehearsal strategies more often and peer learning strategies least often. On the other hand, students attending Anatolian high schools reported using time and study environment, and effort regulation strategies more frequently; and peer learning strategies least frequently.

Table 2. Descriptive Statistics for School Types.

Subscale	Mean		Std Deviation	
	Regular High Schools	Anatolian High Schools	Regular High Schools	Anatolian High Schools
Rehearsal	5.11	4.29	1.26	1.40
Elaboration	5.12	4.51	1.15	1.16
Organization	4.99	4.28	1.24	1.45
Critical thinking	4.67	4.11	1.29	1.28
Metacognitive self-regulation	5.14	4.62	.95	1.00
Time and study environment	5.01	4.82	.95	1.11
Effort regulation	4.98	4.92	1.12	1.23
Peer learning	4.17	3.90	1.41	1.29
Help seeking	4.83	4.63	1.28	1.10

Before conducting MANOVA, the assumptions were checked. It was found that all the assumptions were met. Results of MANOVA revealed statistically significant differences between school types on combined dependent variables ($F(342,9)= 4,662$, $p<0.05$, $\eta^2=.11$; Wilks' $\Lambda= .891$). The multivariate $\eta 2$ based on Wilk's Λ (0.11) indicated that the magnitude of the difference between groups were at moderate level (Cohen, 1988). It implies that, 11% of multivariate variance of the dependent variables was associated with school type. As a follow-up analysis, univariate ANOVAs were run in order to examine the effects of school type on each dependent variable. Bonferroni adjustment was utilized and obtained F statistics were evaluated at the .006 alpha level. Results revealed statistically significant differences between two school types in favor of regular high schools on the cognitive and metacognitive strategies (rehearsal, elaboration, organization, critical thinking, metacognitive self regulation); but not management of different sources (time and study environment, effort regulation, peer learning, and help seeking).

Conclusions and Implications

In the present study, whether students attending Anatolian high schools and regular high schools differ in their use of self-regulatory learning strategies in chemistry were investigated. Results revealed that students attending regular high schools use cognitive and metacognitive strategies (rehearsal, elaboration, organization, critical thinking, and metacognitive self regulation) more frequently. Because students should pass an examination to attend Anatolian high schools, the students in Anatolian high schools are expected to have higher achievement than those in regular high schools. Accordingly, they are expected to use self-regulatory strategies more frequently. Therefore, the result of this study is surprising. The findings of the present study contradict with the results of previous research studies, indicating high achievers use self-regulatory strategies more often (Pintrich & DeGroot, 1990; Pape & Wang, 2003; Yumusak, Sungur, & Cakiroglu, 2007). However, because this study was conducted based on students' responses to self-report questionnaire, deeper investigation is needed to understand when and how students use such strategies by employing qualitative methodologies.

This study focused more on the effects of school type on strategies used in cognitive/ metacognitive, physical and social environment domains in 10^{th} grade level and chemistry course. Future studies should also examine how personal characteristics such as age and gender, or classroom context influence use of these strategies. In addition, other dimensions of SRL such as motivation should be studied as well.

References

Bandura, A. (1986). Social foundations of thought and action: A social cognitive theory. Englewood Cliffs, NJ: Prentice Hall.

Cohen. J. (1988). Statistical power analysis for the behavioral sciences. Hillsdale, NJ: Erlbaum.

Jöreskog, K., & Sörbom, D. (1993). Structural equation modeling with the SIMPLIS command language. Hillsdale, NJ: Erlbaum.

Pape, S. J., & Wang, C. (2003). Middle school children's strategic behavior: Classification and relation to academic achievement and mathematical problem solving. *Instructional Science*, 31, 419-449.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82,* 33-40.

Pintrich, P.R., Smith, D.A.F., Garcia, T., & McKeachie, W.J. (1991). A Manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ). Ann Arbor, MI: National Center for Research to Improve Postsecondary Teaching and Learning., The University of Michigan.

Sungur, S. (2004). An Implementation of Problem based Learning in High School Biology Courses (Doctoral dissertation, The Middle East Technical University, 2004).

Yumusak, N., Sungur, S. & Cakiroglu, J. (2007). Turkish high school students' biology achievement in relation to academic self-regulation. *Educational Research and Evaluation*, 13(1), 53 – 69.

Zimmerman, B. J. (1990). Self-Regulated Learning and Academic Achievement: An overview. *Educational Psychologist,* 25(1), 3-17.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. Pintrich, & M. Ziedner (Eds.), Handbook of self-regulation (pp. 13-39). Orlando, FL: Academic Press.

Zimmerman, B. J. (2001). Theories of self-regulated learning and academic achievement: an overview and analysis. In B. J. Zimmerman & D. H. Schunk (Eds.), Second Edition. Self-regulated learning and academic achievement: theoretical perspectives. (pp. 1-37).Hillsdale, NJ: Lawrence Erlbaum Associates.

Zimmerman, B. J., & Martinez-Pons, M., (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. American Educational Research Journal, 23, 614-628.

SCIENCE EDUCATION RESEARCH AND ITS IMPACT IN THE SCHOOL SCIENCE: LAST DECADES' IN-SERVICE TEACHERS MEMORIES

Roberto Nardi
Universidade Estadual Paulista, São Paulo, Brazil

Maria José P.M. de Almeida
Universidade de Campinas, São Paulo, Brazil

Abstract

Science Education became a solid area of research in the last decades in Brazil, and counts today with important research groups, emerged from de 1970's. These groups are congregated into communities which constitute today what we have called the Science Education or Education in Science area. In a research recently developed about this theme, called "Science Education area constitution in Brazil: memories of former researchers", the authors (NARDI, 2005; NARDI; ALMEIDA, 2004, 2007) interviewed former researchers, considered by theirs pairs as important in this research area constitution. The researchers' discourses interpretation seems to show that there are doubts about the impact of the knowledge accumulated in the last decades in science classrooms. That means: does the knowledge produced by the science education research reached the classrooms? To answer this question we have interviewed basic education science, physics, chemistry and biology teachers who taught or are teaching in the public system of education in Sao Paulo State, in Brazil, in the last few decades, and which are not directly involved with university science education research groups. In this communication we detach outcomes of this research, more precisely three teachers' discourses, through techniques and procedures of discourse analysis, in its French line.

Introduction

Science Education became a solid area of research in the last decades in Brazil, and counts today with solid research groups, emerged from de 1970's of last century and which are responsible for the rise of journals and specific events on physics, chemistry, biology and science teaching, for the development of teaching projects, the organization and coordination of post-graduation programs, in master's and phD levels; in short, these groups are congregated into communities which constitute today what we have called the Science Education or Education in Science area.

In a research recently developed about this theme, called "Science Education area constitution in Brazil: memories of former researchers", the authors (NARDI, 2005; NARDI, ALMEIDA, 2004, 2007) interviewed former researchers, considered by theirs pairs as important in this research area constitution. They opined about factors considered important to the constitution of this area, as well its main characteristics. Interpreting the interviewed researchers discourses we observed that, although they, who contributed for the research area origins, were sure about the importance of the studies produced in Brazil, as well as about the great amount of knowledge accumulated in the area, there are doubts about the impact of this knowledge in the science classrooms, that means: does the knowledge produced in science education has reached the classrooms? This question, put in a more appropriated form, turned into object of a broader research, which we looked to answer through the following question:

How in-service teachers that taught or have been teaching subjects related to the science area, in different teaching levels, and do not belong to the researchers community, have been practicing meanings respect to science education research procedures and results and their possible implications to the teaching they have been practicing?

To answer this question we have interviewed basic education science, physics, chemistry and biology teachers who taught or are teaching in the public system of education in Sao Paulo State, in Brazil, in the last few decades, and which are not directly involved with science education research groups. In this communication we detach partial outcomes of this research, more precisely three of the teachers' discourses, through techniques and procedures of discourse analysis, in its French line.

The methodological and theoretical frames

Teaching discourses have been analyzed using French line discourse analysis principles and procedures, mainly those enunciated by Pêcheux (2002), in France, and Orlandi (1999), in Brazil. According to these authors, language is an interaction among individuals, society, history and ideology and, so, never is transparent; always hold meanings, of different senses, thus, depending on the position the individual enunciates, each word can have different senses. The analyst has to be able to create an analytical device in order to understand how a discourse produces senses which are the discourse production conditions, to establish regularities present and the possible interpretations. So, we detach at least three factors which have to take into consideration in the analysis: the discourse production conditions (the senses relation the discourse has), the anticipation mechanism (which is the capacity the speaker to place himself in the listener's place, in order to think his words effects in somebody that is listening to him – this anticipation will determine the way the individual chooses to speak) and the forces relation (speech's hierarchy), that constitutes a discourse, which is the place the individual speak from (speaking from different places will produce different discourses).

Methods and sample

Data was collected through semi-structured interviews has been taking among a sample of twenty secondary and high school Sciences, Physics, Chemistry and Biology teachers, who taught in the last decades in São Paulo State public schools, in Brazil. The interviews were registered with digital camera, under teachers consent, making clear that the images and the testimonies contents would be used exclusively for research purposes and their names will be kept anonymous. Up to now we have taken ten interviews. The main question refers to the teaching characteristics during their acting period, the didactical resources used, the undergraduate and in-service training the teachers participated. If during the interview the teacher does not give any information about science education research, the interviewer tries to question directly if he/she used research outcomes in their classrooms. In this paper we interpret three of the interviewed teachers' discourses, which were called B (Biology), C (Chemistry) and P (Physics). They taught respectively Biology, Chemistry and Physics in São Paulo State public school in different towns in the last three decades and they have retired recently.

Teachers' discourses

Teachers were first asked by the researcher about the characteristics of their classes and the teaching methodologies they usually performed. The intention was to verify if the teacher interviewed touched in activities resulting from the science education research. We show bellow some examples of teachers' discourses and some of the analysis done.

We perceived, for example, in their speeches characteristics of their former work, such as: the main didactical resources origin in that period was the textbooks. Analyzing the Biology teacher speech, for example, we realized the role textbooks had in that period, like, at the end of the 1960: books and materials.

> *... I took the Biological Science [Biological Science Committee Study] training in São Paulo… So, I used to work with practical work in my classes… […]… that experiment we use to work with the students... three paper strips that, in the first one had printed two kinds of paws, reminding a cock, a chicken, a duck… […] students used to formulate technical hypothesis... I used to work it during all 37 years of my [teaching] life to wake the curiosity, the scientific spirit, and to show that science does not allow guessing; science is based in facts, in experimentation, isn't it? I took that from there [the BSCS]… I took many training…* (B)

When asked about science education research, however, whether they had knowledge about research in this area, or used its outcomes it in their classes, teachers expressed in this way:

No, no... one thing that... only during the undergraduate program… to search how it was… it was not… which was the lack of interest causes… when we had not access to all this developed technology, right? (B)

... We started in the decade... the end of the 80's… in the 90's, that started to appear a lot of stuff… a lot of stuff arrived at school and the majority [of teachers] watched it… with resistance, because there started that video history, all school had to have one… to understand [the lesson] had to use video… I am thinking at the end of the 80's… we still had no computers… So, there was a guy there that used video to record… he arrived at school, in our meetings, set his TV with video, and we only watched all of that. But there was much junk… and many questions had no answer… (C)

We noticed also that in three of the teachers' discourses it was inevitable the comparison between the teaching quality in that period – the beginning of their teaching career – and at the end (they are today retired). To all of them, teaching quality has been deteriorated:

I saw things like... along thirty and three years I have taught, mainly now at the end… the students come at school and, in the first class, says: "I did not like this teacher… [so] I will not… I will not do anything…". And at the end of the school year we could not fail the student; [even though] the student had not worked at all… (P).

We can interpret in their discourses, like above, a strong criticism to public education policies introduced in the last decades by the governments, which they consider responsible by the deterioration of teaching quality. They all detach specially the so called "continuing progression" policy adopted by the São Paulo State School System, in which schools has to generate mechanisms in order pupils have to be recovered, avoiding students failing. The introduction of a 'bonus' aiming to stimulate to improve teachers' frequency and avoiding students' evasion, is also another of the policies criticized by all of the interviewed teachers:

Thus, you cannot ask anything to him; you asked for homework, he/she does not do it… so what? What can we do? You give him/her a quiz… he does not do it; return it blank… so what? So, nowadays, there is nothing we can charge the student… teacher cannot demand anything… does not matter the grade you give him; at the end, he/she does not fail… (C)

He/she [teachers] receives a kind of 'bonus'; this 'bonus' seems to have an influence on…for example… the students' approval percentage, teachers' frequency, students' frequency… So, indirectly, the government compels you… it is some kind of terror against teachers. So it obliges you…. besides not charging the student, there is some intention to terrorize teachers. So, it really obliges you; you have to approve the pupil… [C]

In discussing the initial teachers training period, the undergraduate school they took, one of the interviewed teachers suggests that universities have to demand more, from the future teachers, in terms of contents knowledge, since some teachers do not have enough basic knowledge about the subjects to be taught. The same teacher understands, however, that the current undergraduate courses designed to prepare future secondary and high school teachers have been improved in relation to pedagogical aspects:

Nowadays, it seems to me that there are some schools that were reformed specifically for the teaching. So, looks like to me that this university influence – The Federal [local University]here, I think is like that – it has a science degree directed more to the teaching… So, looks to me that the knowledge [the future teacher] acquires would be more appropriated for teaching. Because the school, before, you attended three years and, only in the fourth one, you would choose: whether 'licenciatura' [to form teachers] or 'bacharelado'[to form science researchers]. [C]

One of the interviewed teachers detaches, also, as positive in the current teachers' preparation programs, the introduction of mechanisms looking to guarantee an interdisciplinary approach among academic disciplines and the consciousness that outside factors can interfere in student school performance and their relationship in the school.

Conclusions and implications

Through this teacher's sample discourse analysis it was possible to perceive that all of them went through continuing education programs during their life as Secondary or High School teachers. We perceived also the presence of universities and/or government policies reflecting in their continuing education process. Even though, teachers have difficulties to explain what they understand by science education research, showing not to recognize what teaching research means. They are also unanimous in criticizing the public education policies adopted by the state governments in the period they acted as teachers, in the last three or four decades, such as the called 'continuing progression'.

It is important, however, to remind that this analysis took into consideration only three of the ten interviews taken up to now, so it is premature to generalize some of the interpretation drawn from the teachers' discourses. More interviews will be taken I this project, in order to confirm this paper partial conclusions. Moreover, we have to interpret these outcomes taking into consideration the historical, economical etc. background, in order to contextualize teachers discourses in this Brazilian history of science education period.

Also, this research could be extend, afterwards, to other states in Brazil, in order to understand what is going on with other teachers' performance in science classrooms, the role played by other universities and state governments about this subject, so, we can match these outcomes with other realities, having an more complete panorama about the role science education research is playing in teachers performance in science classrooms.

References

KRASILCHIK, M. (1996) Formação de professores e ensino de Ciências: tendências nos anos 90. In: MENEZES, L. C. (Org.). **Formação continuada de professores de Ciências.** Campinas: Autores Associados, p. 135-40.

KRASILCHIK, M. (2000) Reformas e Realidade: o caso do ensino de Ciências. **São Paulo em Perspectiva**, 14 (1) 85-93.

LEMGRUBER, M. C. (1999) **A Educação em ciências físicas e biológicas a partir das teses e dissertações** (1981 a 1995): uma história de sua história. Thesis [phD in Education] Faculdade de Educação. Universidade Federal do Rio de Janeiro.

NARDI, R. (2005) **A Área de ensino de ciências no Brasil**: Fatores que determinam sua constituição e suas características segundo pesquisadores brasileiros. Bauru. Thesis (Livre Docência). Universidade Estadual Paulista, UNESP, Faculdade de Ciências.

NARDI, R. (2003) A educação em ciências, a pesquisa em ensino de ciências e a formação de professores no Brasil. **Tecne, Episteme y Didaxis**, Bogotá, Colômbia, Vol. Extra, p. 19-23.

NARDI, R.; ALMEIDA, M. J. P. M. (2004) Formação da área de Ensino de Ciências: Memórias de pesquisadores no Brasil. **Revista Brasileira de Pesquisa em Educação em Ciências**. Porto Alegre: 4 (1) 9-23.

______________________________ (2007) Investigação em Ensino de Ciências no Brasil segundo pesquisadores da área: alguns fatores que lhe deram origem. **Pro-Posições**, 18 (1) 213-226.

ORLANDI, E. P. (1999) **Análise de Discurso: Princípios e Procedimentos**. Campinas: Pontes Editores.

_______________ (2001) **Discurso e Texto:** Formulações e circulação dos sentidos. Campinas: Pontes.

_______________ (1994) Discurso, imaginário social e conhecimento. **Em Aberto**, 14 (61) 52-59.

_______________(1988) **Discurso e Leitura**. São Paulo:Cortez; Campinas: UNICAMP.

_______________ (2004) **Análise de Discurso: Princípios e Procedimentos.** Campinas: Pontes Editores. 4ª. Ed.

PÊCHEUX, M. (1990) **O discurso: estrutura ou acontecimento**. Campinas: Pontes Editores. Transl: Eni Puccinelli Orlandi. 3rd. ed.

EMERGENT OBSERVATION: OBSERVATIONAL SKILLS OF CHILDREN AGED 1 TO 3 YEARS OF AGE

Jane Johnston
Bishop Grosseteste University College Lincoln, UK.

Abstract

This research explored the skill of observation and how it can be supported in children aged between 15 months and 4 years of age engaged in a play activity. The children were videoed playing with a range of toys including, moving toys, aural toys (which made sounds), operated toys (that involved some operation by the child) and soft toys. The children's play was initially categorized according to affective, social functional and exploratory observations and social interactions analyzed for impact on observation. Further analysis focused on children's responses to aural and moving/ operated toys and types of play. Affective observations predominated. Observation in the youngest children was found to be aural, tactile and solitary, whilst in older children, it was more social and functional. There were no real exploratory observations, leading to scientific enquiry, although a few random functional responses bordered on the exploratory, but needed to be supported by the adults present, through scaffolding and modeling. The research indicates the importance of social interaction in play to encourage more scientific play and observations, with adult support being greater in children under 2 years of age and peer support being greater in children between 2 and 4 years of age.

Introduction and Rationale

Observation as an early skill

Observation is an important generic skill for early years children identified in many theories of education (e.g. Pestalozzi, 1894; Piaget, 1929) and subsequently early years curricula (e.g. MOE, 1996; DCSF, 2008) as well as an essential process skill in early scientific development (Covill & Pattie, 2002; de Bóo, 2006). Observation is integral to the development of curiosity and motivation by helping children to remember their investigations. It also helps children to solve investigative problems (Grambo, 1994) and develop process skills (Johnston, 2005). As children develop they begin to focus on specifics in their observation (Harlen & Symington, 1985) and their observations are influenced by personal and taught ideas (Driver, 1983; Duschl, 2000; Tompkins & Tunnicliffe, 2001) as well as interests (Tunnicliffe, & Litson, 2002). Most of our knowledge about observation comes from older primary aged and secondary aged children, although the researcher has some evidence of the nature of observation in children aged 4 to 11 years of age (Johnston, 2009a). This has indicated that observation in young children is tactile involving the sense of touch and hearing as much as sight, but that as children develop they move from broad observations to more specific observations. This research (Johnston, 2009a) also indicates that children bring their previous knowledge to their observations (see also, Duschl, 2000), making fewer theoretical inferences as compared to older children (National Research Council of the National Academies, 2007). The researcher has also found that in very young children, it is often the child's intuitive ideas that take precedence over scientific theory (Johnston, 2005).

Although early years' science is an under-researched area, there is increasing evidence about young children's scientific ideas (e.g. Reiss, M. & Tunnicliffe , 2002; Fleer, 2007; National Research Council of the National Academies, 2007). Scientific skills are less well researched, although the link between observation and understandings is established (Driver, 1983; Duschl, 2000; Tompkins & Tunnicliffe, 2001). Johnson & Tunnicliffe, (2000) in looking at children's understandings of the features of plants, during a visit to a garden and Kameza & Konstantinos, (2006) in researching children's ideas about astronomy found that metacognition and social construction (Shayer & Adey, 2002) were more important than direct observation in the developing scientific understandings.

How professionals support early scientific development

Although the importance of observation is recognized (see above), it has been found to form only 5% of classroom activities and where it was present was not active observation conducted by children, but passive observation made by the teacher (Kallery and Psillos, 2002). It also tends not to be used to initiate activities and motivate children to want to make inquiries (National Research Council of the National Academies, 2007). Practical observational experiences are known to be more appropriate for younger children (BERA, 2003; Howe and Davies, 2005), although the development of good observational skills needs to be supported by focused and structured teaching (Harlen, 2000; Johnston, 2005; de Bóo, 2006), in order to develop thinking and linguistic skills (de Bóo, 2006) and creative thinking (Johnston, 2009b). Pedagogical factors affecting the quality of observational development are similar throughout early years and primary education (Harlen, 2000; Johnston, 2005) and include:

- Time to observe and discuss observations, including the creation of conceptual conflicts (Hand, 1988), through debate and argument (Naylor et. Al., 2004).
- The careful use of observational aids (Harlen, 2000) as they can detract from the observations (Johnston, 2005).
- Encouragement and support from teachers, including questioning (Vygotsky, 1978).
- Focus on patterns, sequences of events and interpretations.
- The use of motivating scientific phenomena or objects to help children to make close observations (Ashbrook, 2007).
- Recording observations. Indeed, rapid sketching of detail has been found (Grambo, 1994) to improve observational skills by focusing on important features which are then remembered.
- The opportunities for explorations where children use their senses, noticing details, sorting, grouping and classifying or sequencing (Johnston, 2005).
- Natural contexts involving the observation of natural phenomena, particularly observations involving animals (Tompkins & Tunnicliffe, 2007).

Without support children are likely to move from their limited unsophisticated creative and imaginative general observations (Tunnicliffe & Litson, 2002) to unsophisticated particular observations, rather than improve their skills in both types of observation. Whilst young children can make very sophisticated and detailed observations, they can get distracted easily and may need support to refocus (Keogh & Naylor, 2003). There has been increased understanding of the pedagogies that support early scientific learning (Harlen, 2000; Kallery and Psillos, 2002; BERA, 2003; Howe and Davies, 2005; Johnston, 2005; National Research Council of the National Academies, 2007; Fleer, 2007), with social interaction being a shared characteristic (Vygotsky, 1962), especially where it involves practical exploration, that builds upon previous knowledge (Piaget, 1929). Active social participation in scientific development appears to be most effective (Johnston, 2009a) with children learning alongside peers (Bruner, 1991) and teachers (Stone, 1993), in a complex social interaction identified by Rogoff (1995). This involves children learning through social interaction on three "inseparable, mutually constituting planes" (Rogoff, 1995 p.139); personal, interpersonal and community/ contextual, which have been found to be useful in analysing early scientific development (Fleer, 2002; Robbins, 2005).

Methods

Having previously identified some aspects of observation in children aged 4 to 11 years of age (Johnston, 2009a), the researcher decided to focus on what the skill of observation looked like in younger children (up to 4 years of age) and how the development skill of observation can be supported in this age group.

The research design draws upon interpretative studies in science education (Lemke, 2001) and in early year's science contexts (Robbins, 2005; Fleer, 2002). Both Robbins (2005) and Fleer (2002) drew upon the analytical techniques of Rogoff (1995) in analysing different aspects of interaction. In this research the individual, peer and adult/ teacher interaction has been analysed in an attempt to understand the skill of observation from both child and adult perspective and the part played by different types of interaction.

The sample involved two groups of children aged between 15 months and 4 years of age engaged in a play activity, where they looked at a range of toys. In one group there were six children aged 15 months to 2 years of age, with 2 early years' professionals, plus the researcher and in the other group there were nine children aged 2 to 4 years of age with four professionals and the researcher. The children attended a private day nursery in a rural location, who had agreed to support the research. All children at the nursery, whose parents had given permission, were included. For each group of children, a collection of toys was placed on the floor. The toys included:

- Moving toys, such as a battery operated hen, which danced while singing, wind-up toys and pull-back cars/ helicopter;
- Aural toys which made sounds, such as a rattle, a battery operated chick which cheeped, a megaphone that children could speak through in alien/ robot/ spacemen voices and a jack-in-the-box;
- Operated toys that involved some operation by the child, such as a ball and hammer set, a wooden frog that makes a frog noise when a stick is pulled across its back, a helicopter (whose propellers move when pushed), a honey bee who 'buzzed' down a pole and colour change ducks (which change colour when warm).
- Soft toys, such as a large dog, a sheep rug (that can be worn).
- Other toys, such as a large multifaceted mirror, a magnetic elephant, with body parts that could be removed and replaced and a wooden person (with moveable limbs).

The play was videoed and although the camera was in full view of the children, only one child took any notice of it during the play and it did not appear to have any effect on the results. The free play activity was observed for 10 minutes whilst the children played independently with the professionals interacting with the children, as appropriate, by playing with a toy, pointing out things the toys did to children and asking them questions about the toy, such as, "Why do you think that happens?"

The video of the interactions was transcribed and analytical induction (Erickson, 1998) was used to identify the types of initial observations made by the children, the number and types of observations made in the different parts of the activity. Further analysis looked at the effect that personal, adult participatory and peer participatory interaction had on the scientific skill of observation (Rogoff, 1995).

Results

The initial observations were grouped into four categories, as in previous research with older children (Johnston, 2009a).:

- affective, showing interest and motivation, such as expressions of glee, "Wow", squeals of delight and giggles;
- functional, observing how the toys work;
- social, involving interactions between children and the adult, such as negotiation for the use of a toy, playing with, or helping another child;
- exploratory, leading to further scientific exploration and inquiry, such as questions that can lead to further exploration or inquiry.

However no children engaged in real exploration, although there was some.professional-directed focus on changes in colour to the colour-change ducks and so this category was dropped from Tables 1a and 1b.

Table 1a: Table to show Affective, Social and Functional Responses to Toys in Children Aged 15 Months to 2 Years of Age.

Number in brackets denotes number of children responding in this way if more than 1

Affective	Social	Functional
Dances (2)	**All are with the professional unless indicated**	Looks closely at elephant
Dances with chicken (2)	Strokes dog	Looks closely at the way a clatterpillar (from the setting's toy collection) moves
Backs away	Looks at Rattle	
Rubs head	Counts ducks	
Waves arms	Takes Jack-in-the-Box to professional	
Cries	Hands ducks to professional one by one	
	Gives dog to professional	
	Brings chicken to professional	
	Professional discusses mirror	
	Professional suggests putting ducks in water	
	Child tries to take a duck from another.	
	Professional encourages child to share ducks	
	Takes chicken to professional	
	Professional focuses on colour change ducks when put in water	
	Professional encourages children to squeeze water out of duck	

Most responses in the children under 4 years of age were no-verbal. Affective responses took the form of dancing and squealing, although a few children used gestures or rubbed their face or heads when unhappy or distressed. Their play and observation was solitary (playing alone) and needed greater adult modeling of play or participation in play. Indeed, the youngest children had no verbal social interaction with peers and limited other social interaction, as can be seen from the interaction below between two of the youngest children aged between 15 months and 2 years of age.

Boy 1 crouches by the hammering box and picks up the hammer and puts in mouth.

Another child comes up to the box and picks it up.

Boy 1 tries to hammer (unsuccessfully, as his motor skills are undeveloped), but continues to play alongside the other child.

The second child takes the box away and Boy 1 picks up the moveable man, drops it and follows the child with the box (with the hammer still in his hand). He drops the hammer and picks it up again.

Table 1b: Table to show Affective, Social and Functional Responses to Toys in Children Aged 3 to 4 Years of Age.

Number in brackets denotes number of children responding in this way if more than 1

Affective	Social	Functional
Dances with chicken (2)	Shows 3 toys to professional without playing with them	Looks closely at rattle (3) while turning moving parts (1)
Squeals with delight (2)	Shows megaphone to a peer	Looks closely at moveable person (3)
Claps hands and giggles and wriggles bottom to sound of chick	Takes dog to peer	Looks closely at rattle (2)
Dances and claps hands and wiggles bottom	Follows another child around and asking for chicken	Looks closely at megaphone (2)
Shows signs of distress when can't have toy	Holds up chicken so another child can't reach it	Looks closely at butterfly (watching it without touching)
Says 'No' when someone takes toy.	Dances with others	Looks closely at kangaroo (observing it moving down a ramp)
Muttering and not looking pleased.	Plays together with dog and sheep	Looks closely at ball (rolling it down a ramp and watching the professional catching it)
Laughs and dances to chicken.	Takes toys to the professional	Looking at the Jack-in-the-box (professional asks, *How can you wind it up?)*
.	Plays builders with other child (for brief moment)	
	Follows other children around	

The older group of children (2 to 4 years of age) occasionally made a verbal response, such as one child who played quietly and independently with the helicopter throughout the whole time observed and then forcefully said *'No!'* when he left it and another child picked it up. Social responses in the youngest age group (15 months to 2 years of age) were initiated by the professionals, who interacted with the toys, *'Stroke the doggy, stroke the doggy',* and gave encouragement to share toys and guidance on the functionality of the toy, *'Push the button'* (to show the child how to turn on the dancing chicken toy). Case Study 1 shows the social interaction involving one girl in the youngest age group (15 months to 2 years). In this interaction, the child initiated some interaction by taking toys to the professional and the professional responded by focusing attention on how a toy works (demonstrating how the Jack-in-the-Box works), asking questions and engaging in ludic play (playing with the ducks and singing *'Four little ducks went swimming one day….'*

Older children (aged 2 to 4 years of age) initially focused on a broader range of toys and engaged in some social responses with peers (see Case Study 2), showing them how a toy worked and sharing a toy with, or taking a toy to, another child, although most of their play was solitary. This can be seen in Case Study 3 where a boy appeared aimlessly to pick up one toy after another, and had almost peripheral engagement with others, or engaged in parallel or companionship play (Bruce, 2004). These Case Studies (1 and 2) also show evidence of Rogoff's, (1995); personal, interpersonal and community/ contextual planes. Some adult interaction, in both age groups, involved questions that focused on the function of the toys, such as, *'What colour has yours gone?'.*(when looking at colour-change ducks with the youngest children, as in case Study 1) *'How do you get that one to work?' 'Push it in'* (when encouraging a child to turn the dancing chicken on himself). These older children needed less adult support to initiate observations, although they appeared to benefit from interaction that focused on specific functions or aspects of the toys. One boy (see Case Study 3) engaged in the most functional responses, by looking very intently at different toys (a rattle, the butterfly, the dancing chicken), but self-initiated functional responses were not particularly characteristic of the children observed.

Case Study 1: To Show Social Interaction involving a Child aged 15 months to 2 years.

0-5 minutes
Girl 1 takes the rattle to professional 2.
She counts the ducks with professional 2, with the professional, singing *'Four little ducks went swimming one day….'*
Girl 1 plays with the Jack-in-the-box with another girl. Professional 2 demonstrates how it works.
The researcher introduces the dancing chicken and Girl 2 dances with some ducks in her hand.
She sits down by the chicken which moves towards her and starts to cry.
Girl 1 goes to the professional1 with the ducks and gives them to her 1 by 1.
Girl 2 moves away from the chicken, still with the ducks in her hand and sits on the cushion with the ducks. She bounces on the cushion in time to the music (still with the ducks in her hand) and stays there for a while.

5-10 minutes
Girl 2 is still on the cushion with the ducks in her hand. She now has all four ducks. She drops two and picks them back up again.
Girl 2 stands up and takes the ducks to professional 2.
A boy comes to play with the ducks and professional 2 encourages Girl 2 to share ducks with him and to put in water and see what happens.
Girl 2 gives the boy two ducks.
Girl 2 looks in the mirror.
She then looks across the room.
She touches mirror.
Girl sits down and plays with the Jack-in-the Box (still with two ducks).
Girl 2 goes to the other side of the room with the ducks and moves back to professional 2.
A third professional brings in a bowl of water and professional 2 encourages Girl 2 to put the ducks in the water.
Girl 2 moves to the water bowl and prods the eater with a duck. She looks at the ducks in the water and professional 2 asks what colour her duck has gone, *'its gone pink'*.
Girl 2 squeezes water out of the duck over the bowl. She holds up high above bowl and squeezes.

Case Study 2: To Show Social Interaction in children aged 2 to 4 years

0-5 minutes
Girl 3 blows into megaphone and shows to Boy 3
Girl 3 shows Girl 4 the megaphone.
Researcher introduces dancing chicken to children and Girl 4 claps her hands and giggles. She then wiggles her bottom to the sound of the music. Girl 4 picks up the chick and takes it away.
Girl 3 follows Girl 4 around, but Girl 4 hold the chicken out of reach. Girl 3 shows signs of distress, putting her hands over her face. She turns away and then goes back to Girl 4. Girl 4 points to another toy and does not let go of the chicken.
Girls 1 and 2 both want chicken, but Girl 4 holds it up. Girl 3 asks for chicken. Girl 3 follows Girl 4 around but Girl 4 holds up chicken out of reach. The professional tells Girl 4 to put the chicken down and watch it dance. She does and then Girl 4 dances and Girl 3 jumps up and down, clapping hands and wiggly bottom.

5-10 minutes
Girl 4 has chicken and Girl 3 follows
Girl 3 puts box on top of chicken, whilst other children watch.
Girl 3 picks up chicken and dances with it.
Girl 3 gets on the floor and shuffles behind the chicken.
Girl 4 lies on the floor and watches Girl 3 and chicken.
Girl 3 squeals with delight.

Girl 4 takes dog and puts it on her head. She takes it to the professional (hanging over her shoulder).
Girls 3 and 4 are playing out of sight with the dog and chicken.

Girl 4, still with the dog, picks up a ball and a duck and takes them to the professional. She wanders away with the dog.

Girl 4 takes the ball to the researcher and rolls it down the ramp.

Case Study 3 – To Show Social and Functional Responses in a Child Aged Between 2 to 4 years of Age

0-5 minutes
Girl 3 shows Boy 2 the megaphone.
Boy 2 watches others and the researcher when she turns on the dancing chicken. He smiles at the chicken.
Boy 2 picks up a colour-change duck
He moves to watch two girls playing with the chicken.
Boy 2 picks up the rattle and looks closely, trying to see how it works.
He helps another child with the hammering (but no real interaction – parallel play)
Boy 2 uses the megaphone when put down by another child, but gives to another child when asked.
Boy 2 picks up the hammer.
Boy 2 plays (half-heartedly) with kangaroo, moving it across the floor.
He plays with the Jack-in-the-Box.
He picks up the rattle and looks closely at it again.
He picks up the moveable man and looks closely at how it works.
Boy 2 dances briefly to the tune from the chicken.
He picks up a colour-change duck.

5-10 minutes
Boy 2 plays with the rattle, again looking at the mechanism.
Boy 2 plays at hammering.
He rolls the rattle over the floor.
He plays with the kangaroo.
The researcher introduces a butterfly and he watches intently.
Boy 2 continues hammering.
He moves a colour-change duck across the floor and flicks it out of the way.
Boy 2 goes to the other end of the room and wanders around.
He plays with another child and puts on a builders hat briefly.
He watches the other child with a builder's hat.
Boy 2 comes back to centre of room and watches a boy with the megaphone and then wanders back and forth, muttering to himself and not looking very pleased.
He laughs and dances to chicken.
Boy 2 follows the chicken. The professional asks if he and another boy want to look at the chicken.
He plays with the chicken.

All children responded to aural and moving/ operated toys, showing interest and motivation through oral exclamations and laughter and by dancing. A few older children also engaged in ludic (fantasy) play and symbolic play (Piaget, 1976) and so the play was analysed again to consider the responses to the aural and moving/ operated toys and ludic and symbolic play (see Table 2). There was overlap between aural and moving/ operated toys, with the toy responded to most often (and receiving the most affective responses) being the dancing chicken, which was both aural and moving. There was also overlap between ludic (fantasy) play and symbolic (epistemic play, which involves children using their existing knowledge about toys in their play. In Case Study 1, Girl 1 (aged between 15 months and 2 years of age) and the professional play with the ducks singing a version of the children's counting rhyme *'Four little ducks went swimming one day….'* In Case Study 2, Girls 3 and 4 (aged between 2 and 4 years of age) play with the toy dog and the chicken and in Case Study 3, Boy 2 used the setting's builder's hat briefly, but does not connect it to the hammering toy. Another child, Boy 2, engaged in both ludic and symbolic play, becoming a sheep by wearing the sheep rug and playing with the dog *'a sheep dog'*, finding a box as a kennel for it. He also, later in the session, used the megaphone, showing it to the professional who encouraged ludic play with the response *'You sound like a robot. Are you a robot?'*

Table 2: Table to show Responses to Aural and Moving Toys and Types of Play (Ludic and Symbolic) at Different Ages.

Number in brackets denotes number of children responding in this way if more than 1

Age	Aural	Moving/ Operated	Ludic (fantasy)	Symbolic (epistemic)
15 months to 2 years of age	Chicken (in constant use) Rattle Jack-in-the-Box	Chicken (in constant use) Hammering balls (4) Rolling balls (2) Rattle Jack-in-the-Box		Stroking dog (2) Ducks (*quack quack* led by professional) Stroking sheep Hammering (picks up hammer and acts out hammering)
2 to 4 years of age	Megaphone (in constant use) for 10 minutes Chicken (in constant use) Jack-in-the-box	Chicken (in constant use) Hammering balls (3) Helicopter (2) Jack-in-the-box (2) Kangaroo Butterfly	Being a sheep (4) Playing with the dog/ being a dog (4) Being a builder (builder's hat) (2)	Dog (making kennel, sheep dog) (2) Playing with sheep Being a builder

Conclusions and Implications

The research appeared to indicate that in younger children, aural and movement act as stimuli to encourage observations, as opposed to the more tactile stimuli, seen in previous research in children over 4 years of age (Johnston, 2009a). As children develop, they engage in more social and functional observations, as well as a few close visual observations. Functional observations appear to require some knowledge not present in very young children to enable them to make theoretical inferences (National Research Council of the National Academies, 2007) and to move to more exploratory observation; that is observation that moves the children to further inquiry. There were no real exploratory observations, although a few random functional responses bordered on the exploratory in that they had the potential to lead to further inquiry. However, in order for this to occur, adult support was needed to encourage the potential in functional observations to move towards more exploratory observations and inquiry.

In the youngest children constant adult support took the form of oral scaffolding and modeling, with the adult playing alongside the child (Stone, 1993), focusing the child's attention on some scientific aspects of the toys and supporting language development (Vygotsky, 1962). With the older children the adult support was partial; the professionals watching the children and with interaction occurring when instigated by the children, or when thought to be socially or pedagogically appropriate. However, this approach led to some missed opportunities, such as in Case Study 3, when the quieter child was not supported in developing their initial functional observations. It appeared that the balance of adult, peer and contextual support was different for different ages, with contextual support being equal in all age groups and adult support being greater in children under 2 years of age and peer support being greater in older children. This balance appears to change again in children over 4 years of age (Johnston, 2009a), with children exercising more autonomy and using prior knowledge in their observations (Duschl, 2000).

The research findings appear to indicate the importance of social interaction in play, encouraging more scientific play and observations This social interaction enables children to negotiate social boundaries (Broadhead, 2004) and develop conceptual understandings through cultural mediation (Bruner, 1991). This confirms ideas concerning effective pedagogy for young children as including interaction between children, their environment and adults (Vygotsky, 1962). Children should be active participants in their own understanding of the world exercising some autonomy and developing understanding from experiences which build upon their previous knowledge (Piaget, 1929). They should have opportunities to scaffold their own and others' learning (Bruner, 1977) with adult support (Stone, 1993). However, it is unclear if this is a conscious pedagogical approach adopted by professionals working with young children. It may be that this needs to be explored more fully with professionals working with very young children to ensure that they move seamlessly from solitary and ad hoc observations, to more socially supported functional and exploratory observations.

References

Ashbrook, P. (2007). Science tickets. *Science and Children,* 45(2), 18.

Australian Academy of Science (2005). *Primary connections*. Canberra: Australian Academy of Science.

BERA, (2003). *Early Years Research: Pedagogy, Curriculum and Adult Roles, Training and Professionalism* Southwell, Notts:BERA.

Broadhead, P. (2004). *Early Years Play and Learning: Developing Social Skills and Cooperation* London: RoutledgeFalmer.

Bruce, T. (2004). *Developing Learning in Early Childhood.* London: Hodder & Stoughton.

Bruner, J. S.(1977). *The Process of Education* 2nd Edition Cambridge, Mass.: Harvard University Press.

Bruner, J. S. (1991). 'The Narrative Construction of Reality' *Critical Inquiry*, Vol.18 No.1: 1-21.

Covill, M. & Pattie, I. (2002). Science skills – The building blocks. *Investigating*, 18(4), 27-30.

de Bóo, M. (2006). Science in the early years. In Harlen, W. (ed) *ASE guide to primary science education* (pp. 124-132). Hatfield: ASE.

DCSF, (2008). *'The Early Years Foundation Stage; Setting the standard for learning, development and care for children from birth to five; Practice Guidance'.* London: DCSF.

Driver, R. (1983). *The pupil as a scientist.* Milton Keynes: Open University Press.

Duschl, R. (2000). Making the nature of science explicit. In Millar, R., Leach, J. & Osborne, J. (Eds). *Improving science education: The contribution of research* (pp. 187-206). Buckingham: Open University Press.

Erickson, F. (1998). Qualitative research methods for science education. In B. J. Fraser & K. G. Tobin (Eds). International handbook of science education (pp. 1155-1173). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Fleer, M. (2002). Sociocultural assessment in early years education: Myth or reality? *International Journal of Early Years Education,* 10(2), 105-120.

Fleer, M. (2007). *Young children: Thinking about the scientific world.* Watson, ACT; Early Childhood Australia.

Grambo, G. (1994). The art and science of observation. *Gifted Child Today Magazine*, 17(3), 32-33.

Hand, B. (1988). Is conceptual conflict a viable teaching strategy?: the students' viewpoint. *Australian Science Teachers Journal*, November 1988, Vol. 34, No. 4: 22-26.

Harlen, W. (2000). The Teaching of Science in Primary Schools 3rd Edition London: David Fulton.

Harlen, W. & Symington, D. (1985). Helping children to observe. In Harlen, W. *Primary science: Taking the plunge* (pp.21-35). London: Heinemann.

Howe, A. & Davies, D. (2005). Science and play. In Moyles, J. (2005). *The excellence of play 2nd Edition* (pp. 154-169). Maidenhead: Open University Press.

Johnson, S. & Tunnicliffe, S. D. (2000). Primary children talk about plants in the garden. P*aper presented to the NARST Conference, April 2000, USA.*

Johnston, J. (2005). *Early Explorations in Science 2nd Edition.* Maidenhead: Open University Press.

Johnston, J. (2009a). What Does the Skill of Observation Look Like in Young Children? To be published in *International Journal of Science Education.*

Johnston, J. (2009b). 'Creativity in Science Education' in Wilson, A. (2009) *Creativity in Primary Education 2nd Edition.* Exeter: Learning Matters.

Kallery, M. & Psillos, D. (2002). What happens in the early years science classroom? The reality of teachers' curriculum implementation activities. *European Early Childhood Education Research Journal,* 10(2), 49-61.

Kameza, M. & Konstantinos, R. (2006). An approach to the introduction of elementary astronomy concepts in early education. Paper presented at the *European Conference on Educational Research, University of Geneva 13-15 Sept 2006*.

Keogh, B. & Naylor, S. (2003). 'Do do as I do'; being a role model in the early years. *Primary Science Review.* No. 78: 2003.

Lemke, J. L. (2001). Articulating communities: Sociocultural perspectives on science education. *Journal of Research in Science Teaching,* 38(3), 296-316.

MOE (Ministry of Education) (1996). *Te Whãriki early childhood curriculum*. Wellington, New Zealand: Learning Media.

National Research Council of the National Academies (2007). *Taking science to school.* Washington, DC: The National Academies Press.

Naylor, S., Keogh, B. & Goldsworthy, A. (2004). *Active Assessment: Thinking, Learning and Assessment in Science.* Sandbach, Cheshire: Millgate House.

Pestalozzi, J. H. (1894). *How Gertrude Teaches her Children* translated by Lucy, E. Holland and Frances C. Turner. Edited with an introduction by Ebenezer Cooke. London: Swan Sonnenschein.

Piaget, J. (1929). *The Child's Conception of the World.* New York: Harcourt.

Piaget, J. (1976). 'Mastery Play' and 'Symbolic Play' in Bruner, J., Jolly, A., and Sylva, K. in *Play – Its role in Development and Evolution.* Middlesex: Penguin.

Reiss, M. & Tunnicliffe, S. D. (2002). 'An International Study of Young People's Drawings of What is Inside Themselves' *Journal of Biological Education* Vol.36 No.2: 58-64.

Robbins, J. (2005). 'Brown paper packages'? A sociocultural perspective on young children's ideas in science. *Research in Science Education*, 35(2), 151-172.

Rogoff, B. (1995). Observing sociocultural activity on three planes: Participatory appropriation, guided participation, and apprenticeship. In Wertsch, J. V., Del Rio, P & Alvarex, A, (Eds.), Shayer, M. and Adey, P. (2002). (Eds.) *Learning intelligence. Cognitive acceleration across the curriculum from 5 to 15 Years.* Buckingham; Open University Press.

Sociocultural studies of mind (pp. 139-164). Cambridge, UK: Cambridge University Press.

Stone, C. A., (1993). 'What is Missing in the Metaphor of Scaffolding?', Forman, E. A., Minick, N., Stone, C. A, (1993) *Contexts for Learning; Sociocultural Dynamics in Children's Development* New York: Oxford University Press: 169-183.

Tomkins, S. and Tunnicliffe, S.D, (2007). Nature tables: Stimulating children's interest in natural objects. *Journal of Biological Education*, 41(4), 150-155.

Tunnicliffe, S. D. & Litson, S. (2002). Observation or imagination? *Primary Science Review,* 71, 25-27.

Vygotsky, L. (1962). *Thought and Language* Cambridge. MA: MIT Press.

Vygotsky, L., Cole, M. (ed) (1978). *Mind in Society, The Development of Higher Psychological Processes.* Cambridge, Mass: Harvard University Press.

THE IMPACT OF PROJECT- BASED LEARNING APPROACH IN SCIENCE EDUCATION ON PRE-SERVICE TEACHERS' ATTITUDES FOR SCIENCE AND PROJECT

Nilgün Yenice
Adnan Menderes Üniversitesi

Abstract

In this study, project-based learning approach was compared with classical learning approach which has been currently utilized in science courses, and the impact of the approach on the attitudes of pre-service teachers about science and project studies was researched in a university in Turkey. The research design was the pre-test/post-test control group. Science Education-I course in Education Faculty was executed with classical approach as control group, and with project-based learning approach as experimental group. The data base of the search was interpreted by the help of SPSS 11.0. T-test was used for analyzing the data. The analyses of the data illustrated that project-based learning approach has a positive impact on the experimental group with regard to science and project studies. Moreover, it was found that gender does not create a significant difference on the attitudes of students regarding science and project studies.

Introduction

The countries which are in front regarding science and technology are using the instructional techniques they have developed in accordance with technology in classroom environments. Individuals are no longer the passive receivers of the information belonging to the world but the active users of it and the creators of the development and change. In this context, educational systems should take on the mission of preparation of the individuals to this kind of world and aim to grow "learner" individuals. The mentioned structure necessitates the approach of student "reaching knowledge, choosing the desired knowledge from complex information networks and solving his/her problems by using this knowledge", instead of the approach of student "receiving the knowledge from teacher" (Akdeniz and Devecioğlu, 2001; Korkmaz and Kaptan, 2002). In other words, comprehensive learning, rather than memorizing, solving problems about a newly introduced situation and development of science process skills are of great importance in today's world. Science lessons are leading among the lessons that students acquire these skills (Kaptan and Bozkurt, 2002).

The attention that is paid on science and technology education can be clearly seen on the educational programs of many developed countries. For instance, content standards of national science program of the United States have eight fields, one of which is science and technology. "Design and Technology" and "Knowledge Technologies" lessons are educated in England, starting from the first class of primary school. "Science and Technology" lessons are taught to students in Japan and Canada (NSES, 1997; Yamazaki and Savage, 1998).

The goal of science education in Turkey is to make students acquire the scientific knowledge, develop their abilities and attitudes (Çepni and Akdeniz, 1996). Elementary level science education is effective for students' recognition of the world, understanding the events and developing the cognitive process skills. Bringing up the elementary teachers well regarding science education is one of the agents that will affect the quality of education to give the mentioned characteristics.

From this point of view, the analysis of "attitudes" which is one of the most important elements of learning activities gets importance. Attitudes are broad learning acts including opinions, values, and assessments executed by individuals. These acts may be either positive or negative.

The project works have a great share on the achievement of the aimed abilities in science. As the students enjoy and have the opportunity of learning by living, they understand the subjects well (Winn, 1997, Lewis et al., 2002).

Project based learning is an instructional and teaching approach focused on concepts and scientific principles including students' problem solving and other significant learning activities, letting students study on their own to construct their own knowledge and providing them a realistic environment to exhibit and end up their studies (Cole, 2002; Railsback, 2002).

In the project based learning perspective, learning is dealt with the reorganization of the cognitive structure of the learner. This understanding stresses the importance of its process intended structure and learning style that enables it as an interactive learning environment in accordance with the changing circumstances. These learning environments are the technology based learning places where the students change and direct their learning themselves to develop their creativity, solve the problems they have in a collaborative manner, give decisions about their success, transfer the real-life experiences into the classroom and include their families actively in the learning process (Erdem, 2002). As Solomon (2003) indicates, in project based learning, students work in groups for solving problems which are multi-disciplinary and related to the program. Trained students make their own decisions about how to approach the problems and what kind of activities to do. They collect data from various sources and provide knowledge by analyzing and synthesizing. Student's independent decision making on how to solve and follow the steps of a problem is the main characteristic of the project (Kubinova, Novotna and Litter, 1998). Project works may be held individually or in groups. There is not any stated lesson hour for the application of the project work. Students may study on their projects in any proper place at any time. Projects, in line with their characteristics, can take place in the daily issues of students (Hamilton and Hamilton, 1997). In the project based learning process, the role of the teacher has evolved (Dopplet, 2003). Besides their classical roles, the teachers are responsible for creating the atmosphere that enables the team spirit of students by letting them learn flawlessly. Teachers direct students to think, use his or her knowledge, and apply their own ideas to their original projects.

Another aim of the project studies is to give students the ability to do scientific research and to provide them "learning by living" (Raghavan et al., 2001). Projects give students opportunity to create something that demonstrates their knowledge of a topic. Typically a project demonstrates cognitive knowledge (Lund ve Tannehill, 2005).

The conducted studies showed that at university level, especially the project based learning approach plays an important role in assessment process (Sezgin, et al., 2001). This kind of project work enables students to have an augmented self-confidence and to be independent learners. In this context, the importance of project studies in science education has risen as a result of this approach (Kaptan, 1999). The students' attitudes towards this approach will make it possible to get information about the functionality of the approach.

In a study conducted about project based learning approach, it is stated that the project works are not emphasized sufficiently in high schools and students lack the desired level of project conduction (Akdeniz and Devecioğlu, 2001). The same statement is valid for the prospective physics teachers in the sense that their project preparation ability is not sufficient (Akdeniz and Keser, 2000). However, it is found that 7th grade students who are given a science education based on project works have an increased academic achievement, efficiency and working hours (Korkmaz and Kaptan, 2002). The same researchers (Kaptan and Korkmaz, 2002) find in their study conducted with the same sample that project based learning affects the creative thinking, problem solving and academic risk taking of the students positively. Project based learning approach is found to have positive impact on 6^{th} grade students in a research which has taken intelligence type inclinations of students into consideration

(Özdener ve Özçoban, 2004). Moreover, it is asserted that a selection of an appropriate teaching method is vital for students' individual interests and talents (Özdener ve Özçoban, 2004). Apart from these studies, Sert-Çıbık (2006) carried out a research with the aim of investigating the effect of project based leaming approach in teaching science to the logical thinking ability and attitude towards science lesson of seventh grade students of primary education. In the research, it is concluded that science education that has been developed with taking the project based learning approach in center.

As a result of literature survey, it is seen that the effectiveness of project-based learning approach in different fields was researched before. In this study, the feasibility of project-based learning as a new understanding in Turkish education system in science teaching was investigated. In the researches done related to the field of project-based learning approach, it's stated that students found this approach useful and practicable. The studies also indicate that the project-based learning approach affect students' motivation towards studying in a positive way. At the same time, with the project based learning approach, students' humanistic feautures that involves capabilities of thinking creatively and critically, having responsibility, making decisions, solving problems, reaching, using and sharing information and making researches which are required by information society, stand out.

Thanks to learning and teaching activities planned and conducted by project based learning approach students construct their own knowledge in case of a certain problem situation by themselves. Students prepare themselves to real life well with project based learning approach. In this study, the feasibility of project based learning approach which can be an innovator approach for educating students equipped with qualities mentioned before was examined in science teaching and designed as a solution for problems in science teaching.

In this context, the necessity of analyzing of pre-service teachers' attitudes towards science and project studies becomes apparent. This study, raised from the mentioned necessity, is hoped to be helpful to science educators, program developers, pre-service teachers.

Purpose

The main purpose of this research was to compare project based learning approach and classical learning approach, applied in science lessons, and to determine the effects of these educational methods on the attitudes of pre-service teachers in education faculties. One hundred third-year pre-service teachers taking the "Science Education-I" course in the Faculty of Education, Adnan Menderes University, Turkey, were chosen as participants for this research. The questions below were aimed to be answered in the light of the outcome of the research:

1. Is there any significant difference between the experimental group that is applied project based learning approach and the control group that is applied traditional learning method regarding their attitudes towards science?
2. Is there any significant difference between the experimental group that is applied project based learning approach and the control group that is applied traditional learning method regarding their attitudes towards project studies in science education?
3. Is there any significant difference between the pre-test and post-test attitude scores of the students towards science and project studies according to gender in science education?

Methods

The model of the research

This study compares project based learning approach and classical learning approach at the same time, searches for the effects of the project based learning approach on the attitudes of the pre-service teachers' towards science and project studies. In addition to that, it is analyzed if there is any significant difference between the pre-test and post-test attitude scores of the students towards science and project studies according to gender in science education. In this manner, experimental design according to "pre-test, post-test" approach with control group is used.

The research group consisted of one hundred third-year pre-service teachers (59% male, 41% female) from the Department of Primary Teaching of Adnan Menderes University. Both experimental and control groups had 50 students. To provide the equality of the groups, the age, gender and weekly science hours were compared. After the statistical analyses, it was determined that the groups were equal. While composing the groups, it was found that the pre-test results of the classes were not significantly different from each other, so the two classes were randomly assigned as experimental and control groups. Both experimental and control group were taught by the researcher.

Lessons in control group were taught by repetition, question-answer techniques and teacher and classbook centred lesson plans which were prepared according to classical learning approach. In control group, teacher centred classical learning approach, in which the teacher is active and the student is passive in class experience and in educational activities in that those experiences are given, was performed.

"Science Education-I" course was instructed via project based learning approach in the experimental group for 15 weeks. At the first phase of the study, the experimental group students were acquainted with project based learning approach, its application steps and benefits. The 50-person experimental group is divided into project groups consisting of 5 people. Each project group chose a representative group name. Project plans were prepared for the groups, learning objectives and project steps were determined in detail by the interviews with each group. Project presentation topics were synchronized with the 4th grade primary school Science curriculum-and its subject content. Students were encouraged to study in line with their interests. Group members distributed the responsibilities on their own and made a working schedule to utilize the designated time most efficiently. Students studied both individually and in groups actively during the project studies. Guidance and help were offered for the groups in project design phase, not only in class environment but also out of the class through office hours. Evaluation forms were used at mid-evaluation by the students and the project executor. Projects presented in the light of developed criteria are evaluated with regard to documentation and application. Group projects were then exhibited at the end-year science festival to participants.

Data collection instruments

In this study, "Attitude Scale Towards Science" developed by Baykul (1990) and "Attitude Scale Towards Project Studies" developed by Sezgin et al. (2001) were used in order to determine the attitudes of the pre-service teachers towards science and project studies. "Attitude Scale Towards Science" consisted of 30 items with Cronbach's Alpha reliability coefficient of 0.94. The items in this scale was composed in 5 point Likert scale ranking from certainly agree, usually agree, no idea or indecisive, disagree to strongly disagree. The Likert type scale, consisting of 21 items with Cronbach's Alpha reliability coefficient of 0.92, determined the attitudes of the students towards project studies. Before applying the scales, the reliability of the scales were assessed by the researcher again. The application was done to 55 first-year pre-service teachers. Cronbach's Alpha reliability coefficient of "Attitude Scale Towards Science" was found .89 and the "Attitude Scale Towards Project Studies" was found .87. In both scales the positive statements were ranged from "certainly agree" - scored as 5,4,3,2,1 - and the negative statements are ranged from "strongly disagree" or "against" - scored as 1,2,3,4,5. The maximum and minimum scores that could be obtained from "Attitude Scale Towards Science" were 150 and 30 respectively; the maximum and minimum scores that could be obtained form "Attitude Scale Towards Project Studies" were 105 and 21 respectively.

Analysis of data

The data gathered from the attitude scales and the personal data form were initially coded, then analyzed by using SPSS 11.0 Statistics program. Dependent t-test was applied in order to determine the difference of the attitude scores of the control and experimental groups. Independent t-test was applied in order to determine the difference of the attitude scores of the students applying project based learning approach towards science and project studies according to gender.

Results

T-test was used to determine if there was any significant difference in the pre-test scores of students of the control and experimental groups at the beginning towards science and project studies (Table 1).

Table 1. The Pre-test scores of students attitude in control and experimental groups towards science and project studies

		N	X	SD	t	P
Attitude scale towards science	Experimental Group	50	104.36	18.93	1.78	0.78
	Control Group	50	97.50	19.59		
Attitude scale towards project studies	Experimental Group	50	75.48	10.23	-0.22	0.82
	Control Group	50	76.02	13.56		

As it can be seen from Table 1, there is no significant difference between the attitudes of the students in the control and experimental groups towards science ($p>0.05$) before starting to teach. It can be observed that the attitudes of students in the control (X= 97,50) and experimental groups (X=104,36) towards science are close to each other. At the same time, it is also observed that the attitudes of the students in the control (X= 76,02) and experimental groups (X=75,48) towards project studies are close to each other at the beginning of the study. In Table 1, there is no significant difference concerning project studies attitudes between the experimental groups and control groups. Thus, it is obvious that the project studies attitude level of the groups has similar properties at the beginning.

Table 2. Pretest - posttest scores of the experimental group students towards science and project studies

		N	X	SD	t	P
Attitude scale towards science	Pre-Test	50	104.36	18.93	-4.41	0.00*
	Post-Test	50	111.62	18.16		
Attitude scale towards project studies	Pre-Test	50	75.48	10.23	-5.89	0.00*
	Post-Test	50	87.36	14.19		

*$p<0.05$

The pre-test and post-test results of the experimental group students from the attitude scale towards science and project studies are illustrated in Table 2. Since $p=0.00<0.05$, there is a significant difference (at 0,05 level), between the pre-test and post-test points of the attitudes towards the science of the experimental group students. This result shows that project based learning approach affirmatively affects the attitudes of the experimental group students towards science. At the same time, there is a significant difference at between the pre-test and post-test scores of the attitudes towards project studies of the experimental group students. We can derive that project based learning approach affects the attitudes of students in the experimental group positively towards the project studies.

Table 3. Pretest -posttest scores of the control group students towards science and project studies

		N	X	SD	t	P
Attitude scale towards science	Pre-Test	50	97.59	19.78	-0.20	0.84
	Post-Test	50	98.00	24.25		
Attitude scale towards project studies	Pre-Test	50	76.02	13.56	-1.83	0.07
	Post-Test	50	79.46	12.50		

When the pre-test and post-test results of the control group students towards science are examined, it can be seen that there is no significant difference between the pre-test and post-test results since $p=0.84>0.05$. This result demonstrates that students' attitude towards science is not affected during research period in the control group where the lessons are taught traditionally.

Clearly indicated in Table 3, it can be seen that there is no significant difference between the pre-test and post-test results since p=0.07>0.05 when the pre-test and post-test results of the control group towards project studies are examined. This outcome shows that students' attitude towards project studies is not affected during research period in the control group where lessons are taught traditionally.

Table 4. Post-test scores of the experimental and control group students towards science and project studies

		N	X	SD	t	P
Attitude scale towards science post-test	Experimental Group	50	111.62	18.16	3.30	0.001*
	Control Group	50	97.44	24.32		
Attitude scale towards project studies post-test	Experimental Group	50	87.36	17.19	2.95	0.004*
	Control Group	50	79.46	12.50		

*p<0.05

As can be observed in Table 4, there is a significant difference between post-test scores of the experimental and the control group students attitudes towards science and project studies.

Table 5. Pre-test and post-test scores of students participating the research towards science for gender

		N	X	SD	t	P
Attitude scale towards science pre-test	Female	64	99.27	20.77	-1.14	0.25
	Male	36	103.89	16.78		
Attitude scale towards science post- test	Female	64	103.59	22.41	-0.75	0.45
	Male	36	107.14	22.40		

In Table 5, it is seen that there is no significant difference of the attitude scale towards science pre-test results between the attitudes of students towards science and their gender when pre-test and post-test results (p=0.25>0.05 for females, p=0.22>0.05 for males) of students are compared. Also, there is no significant difference of post-test scores of the Attitude Scale towards Science (p=0.45>0.05 for females, p=0.45>0.05 for males). It can be concluded that gender is not effective in students' attitude towards science. This finding is also in concordance with the findings of various other researches (Henderson, Fisher ve Fraser, 1998; Boone, 1997; Neathery, 1994; IAEP, 1992).

Table 6. Pretest-posttest scores of students who participate the research towards project studies for gender

		N	X	SD	t	P
Attitude scale towards project studies pre-test	Female	64	76.18			
	Male	36	74.97	10.19	0.48	0.62
Attitude scale towards project studies post-test	Female	64	82.34		-1.02	0.30
	Male	36	85.31	11.67		

In Table 6, it is seen that there is no significant difference when pre-test and post-test results of students' attitudes towards project studies and gender are compared (p=0.62>0.05 for females, p=0.66>0.05 for males). Also, there is no significance of post-test scores of the Attitude Scale towards Project Studies (p=0.30>0.05 for females, p=0.36>0.05 for males). It is observed that gender is not effective in students' attitude towards project studies.

Discussion of Results

The results of this research which was conducted in order to determine the effects of the project studies on the attitudes of the pre-service teachers towards science and project studies are as follows:

1. Before the beginning of the research, it was found that there was no significant difference between the experimental and control groups students' attitudes of science and project studies. Therefore, the levels of both groups were assumed equal at the beginning.
2. The experimental group students have more positive attitudes towards project studies than the control group students. The experimental group students' attitudes are positively affected by the instruction method pursued in project based learning approach.
3. The experimental group students who are educated with project based learning approach have more positive attitudes towards science than the control group students. Project based learning approach affects the students' attitudes towards science.
4. No significant change is observed in the attitudes of control group students towards science and project studies.
5. The general attitudes of students participated in the research do not show any difference in respect to their gender.

It is found with the applications of two different study groups that project based learning approach is more effective than classical approach. It is also found that project studies in the experimental group encourages the active participation of the students and largely contributes the students' learning by doing-trying-applying. It is observed that motivation is positively affected since the students themselves decide the topics of the projects. Students that are divided into groups find the opportunity to determine their own studying styles and to work on the settings they decide. In group studies, students can thoroughly think, solve problems, find and learn how to acquire knowledge and boost their creativity. Results obtained from this study show parallelism with similar prior researches (Boundria, 2002; Cargo, 2000).

Science lessons are composed of many subtopics. In this research, it is found that students can apply their know-how via project based learning approach and can have a more permanent and efficient retention of knowledge on various areas of science. It should be noted that the development levels of the students in the lessons within project scopes other than science lesson (such as social sciences, mathematics etc.) are not investigated in this research. However, it is predicted that enhancement in students' attitudes towards mentioned lessons will be observed in the light of prospective field studies.

Studying in groups lead students to augment their knowledge and experience, to gain the sense of group spirit, to find the chance to come up with new ideas and questions, to share the knowledge and to learn more efficiently. Students say in the interviews that they are content to study with team spirit and to be a part of a project. Many similar researches (e.g.: Curtis, 2002) has consistency with the finding of this research. Yet, it is hard to prove that group members equally contribute to the projects and this reality evokes anxiety among the other members of the group members. These facts are the results of some researches (Elliott and Higgins, 2005). Giving individual projects along with group projects will be helpful for the students that express themselves individually more clearly. On the other hand, some students have hard times to approach with questions, even to speak in classes where interactive instruction method is applied. Working in groups may form beneficial relationships among students and develop their self-esteem.

Conclusions and Implications

Permanent and efficient learning is a target in project based learning approach with active participation of the student. In this context, the project executor has important responsibilities such as preparation of the project plan, determination of sources and tools, continuous supplementation of the project with innovative changes via observation, and control of student activities and knowledge transfers.

In the light of the researches and statistics, we can say that students can learn more permanently and more efficiently by their active participation and this helps them to develop positive attitudes about the lesson. The enhancement in the attitudes will surely trigger the success of the students. İmer (2008) demonstrates the effects of project based learning approach to the students' academical achievement an attitude in elementary science and technology education. According to the results of research, there is a meaningful difference the attitude and suuccess toward Science Education course between the group which has been applied project based learning approach and the group which has been applied traditional learning method. On the other hand, importance of the concept "attitude" ensues. Essentially, the teachers' attitude towards the students is rather crucial. Attwater, 2000 ve Foster, 1997 focus on that positive and encouraging attitude of the teacher contributes substantially to the students' tendency to be involved, do excellent work and present high attention in class (Tal, Krajcik and Blumenfeld, 2006).

The following suggestions are offered depending on the results of this research.

1. This study is a beginning for using project based learning in science education. So, this study can be developed. The sample of the research can be increased, besides attitudes of teachers, their opinions can be received.
2. It was found out that students' attitude towards project studies quite positive. Therefore, the teacher candidates should be directed more project studies.
3. In science courses the lecturing should be avoided, instead project based learning approach should be used with other teaching and learning approaches.
4. Not only for the science lessons but also for the other lessons (mathematics, social sciences, Turkish and etc.) the application of the method should be supplied in Turkey.
5. First of all, the pre-service teachers should be educated, in order for the projects to be conducted at the desired quality and make their students in the future obtain the aimed behaviours. In this context, it is obligatory for pre-service teachers to gain the ability of designing student centred activities. It is thought that, especially the application of researcher teacher growing approach would be useful for this situation.
6. It is also believed that, the in-service teachers' participation to whether the graduate studies in their fields or courses, seminars, conferences and etc. and their completion of studies that make students active and give them the learning responsibility would contribute to the carrying out of the student projects.
7. There is necessity about the substructure of the primary schools regarding laboratory, computer, internet, library and etc. should be improved for the application of the project studies at the students' level.

It is known fact that project based learning approach is not favored by the teachers in Turkey due to time constraints and stress regarding catching up with the topics covered in curriculum. In this respect, utilization of computers and internet will be helpful. Such kinds of intercommunication media provide easier and faster eventuation of project based learning approach. Web sites in other countries developed with this topic (such as www.learner.org; www.thinkquest.org) have to be widespread among teachers. This will increase the acquaintance with project based learning approach.

References

Akdeniz, A.R. & Y. Devecioğlu. (2001). The evaluation of project studies executed in secondary school science lesson, *Maltepe Universtiy Science Symposium*, Istanbul: Maltepe University.

Akdeniz, A. R. & Ö. F. Keser. (2000) An approach to develop the project preparation skills of science teacher candidates, *Science Education Congress IV*, Ankara: Hacettepe University.

Baykul, Y. (1990). *The Changes in the Attitudes regarding Mathematics and Science Lessons from Primary School Fifth Grade to Last Grade of High School and the Factors that are Considered to be Linked with the Success in "Student Entrance Exam for Universities"*, Ankara: OSYM Publications.

Boone, W.J. (1997). Science attitudes of selected middle school students in china: a preliminary investigation of similarities and differences as a function of gender, *School Science and Mathematics, 97*(2).

Boundria, T.J. (2002). Implementing a project-based technology program for high school women, *Community College Journal of Research and Practice, 26*,:709-722.

Cargo, R.A. (2000). Made for each other: Nonprofit management education, online technology, and libraries, *The Journal of Academic Librarianship*, *26*(1): 15-20.

Cole, K., Means, B., Simkins, M. & F. Tavalin. (2002). *Increasing Student Learning Through Multimedia Projects.* Virginia/Alexandria: Association for Supervision and Curriculum Development.

Curtis, D. (2002). The power of project", *Educational Leadership*, *60*(1):50-53.

Çepni, S. & A.R. Akdeniz. (1996). A new approach to train science teachers, *Hacettepe University Education Faculty Journal*, *12:*221-226.

Doppelt, Y. (2003). Implementation and assessment of project based learning in a flexible environment, *International Journal of Technology and Design Education*, *13*:255-272.

Elliott, N. & A. Higgins. (2005) Self and peer assessment-does it make a difference to student group work?, *Nurse Education in Practice*, *5*(1):40-48.

Erdem, M. (2002). Project based learning. *Hacettepe University Education Faculty Journal*, *22*:172-179.

Hamilton, M. A. & S.F. Hamilton. (1997). When is work a learning experience?, *Phi Delta Kappan*, *78*(9):682.

Henderson, D. G., Fisher, D. L. & B. J. Fraser. (1998). Learning environments in senior secondary environmental science classes*, The Annual Meeting of the Australian Association for Research in Education*, Adelaide, South Australia.

Imer, N. (2008) The Investigation of the Effects of Project Based Learning Approach to the Students' Academical Achievement and Attitude in Elementary Science and Technology Education, Post Graduate Thesis, Gazi University, The Institute of Educational Sciences

International Assessment of Educational Progress (IAEP). (1992). *Learning Science* (Princeton, NJ, Educational Testing Service).

Kaptan, F. (1999). *Science Education,* Ankara: Ministry of Education Teacher Books Series.

Kaptan, F. & H. Bozkurt. (2002). Project based learning approach and science festival, *Contemporary Education*, *287*:18-28.

Kaptan, F. & H. Korkmaz. (2002). The effect of project based learning approach in science education on creative thinking, problem solving and academic risk taking, *Hacettepe University Education Faculty Journal*, *22*:164-170.

Korkmaz, H. & F. Kaptan. (2002). The effect of project based learning in science education on academic success, academic personality concept and study hours among primary school students, *Hacettepe University Education Faculty Journal*, *22*:91-97.

Kubinova, M., Novotna, J. & G. H. Littler. (1998). Projects and mathematical puzzles- a tool for development of mathematical thinking, *European Research in Mathematics Education* I, Iı:Group 5.

Lewis, S. P., Alacaci, C., O'Brien, G. E. & J. Zhonghong (2002) Reservice elementary teachers' use of mathematics in a project- based science approach, *School Science & Mathematics,* 102, 4.

Lund, J. L. & Tannehill, D. T. (2005) Standards-based physical education curriculum development (http://books.google.com.tr)

Mayer, R.E. (1998). Learning strategies, in: Claire E. Weinstein, Ernest Goetz, Patricia A. Alexander (Eds.) *An Overview Learning and Study Strategies,* San Diego/California: Academic Pres. Inc.

National Science Education Standards (1997). *National Research Council*, Washington DC..

Neathery, M.F. (1994). Elementary and secondary students' perceptions toward science: correlations with gender, ethnicity, ability, grade, and science achievement, *EJSE*, 2(1).

Özdener, N., T. Özçoban. (2004). The effect of project based learning modal parallel with multiple intelligence theory on students' success in computer education, *Educational Sciences in Theory and Practice*, *4*(1):147-170.

Project Based Learning. (2002). *BIE (Buck Institute for Education),* Retrieved May 18, 2007 from http://www.bie.org.

Raghavan, K., Coken-Regev, S. & S.A. Strobel. (2001). Student outcomes in a local systemic change project, *School Science and Mathematics*, *101*(8).

Railsback, J. (2002). *Project-Based Instruction: Creating excitement for learning planning and program development,* Nort West Regional Educational Laboratory. http://www.nwrel.org/request/2002aug/ project based.

Sert-Çıbık, A. (2006) The effect of project based learning approach to the logical thinking ability and attitude of students in science lesson, Post Graduate Thesis, Çukurova University, The Institute of Social Sciences

Sezgin, G., Çalışkan, S., Çallıca, H. & M. Erol. (2001). Students' attitudes on project based laboratory studies in science education, *Maltepe Universtiy Science Symposium*, Istanbul: Maltepe University.

Solomon, G. (2003). *Project Based Learning: A Primer Technology & Learning.*

Tal, R., Krajcik, J. & Blumenfeld, P. (2006). Urban schools teachers enacting project-based science Retrieved September 16, 2009 from http://www-personal.umich.edu/~krajcik/Papers.htm

Thomas, J. W. *A Review of Research on Project Based Learning*, (2000). Retrieved April 23, 2007 from http://www.autodesk.com/foundation..

Winn, S. (1997). Learning by doing: Teaching research methods through student participation in a commissioned research project, *Studies in Higher Education*, *20* (2).
Wolk, S. (1994). "Project Based Learning: Pursuits with a purpose". *Educational Leadership,* 52(3), 42-45, November.
Yamazaki, S. & E. Savage. (1998). Views of technology education in canada and the united kingdom, *The Journal of Technology Studies*, *14(*1)

RELATIONS BETWEEN MOTIVES, ACADEMIC ACHIEVEMENT AND RETENTION IN THE FIRST YEAR OF A MASTER PROGRAMME IN ENGINEERING PHYSICS

Staffan Andersson
Uppsala University

Cedric Linder
Uppsala University

Abstract

Students' motives for enrolling in a master programme in Engineering Physics at a traditional research university in Sweden were explored using an open-ended questionnaire and James Gee's notion of Discourse Models. Four different Discourse Models of how students explain their reasoning for enrolling in the programme were identified. These Discourse Models have been characterized as Programme student, Engineer to be, Cosmic explorer and Convenience student. They are discussed in relation to student retention and achievement in the programme. The relationship between these Discourse Models and academic achievement during the first year of study was explored by cross-referencing the motivational patterns of students with academic records. This showed a connection between levels of achievements and students' Discourse Models. The group using the programme student model showed the highest fraction of successful students, whereas a majority of the students using the engineer to be model were unsuccessful or left the programme. Students using the convenience student model had the largest fraction of unsuccessful and "leaving" students. This study further explores some of the complexities of dynamics regarding engineering student achievement and attrition. The study exemplifies the usefulness of a Discourse Model perspective when exploring such issues to inform teaching and learning.

Introduction

Internationally, a continuing general decline in interest for pursuing studies in science and engineering programmes continues to challenge many institutes of higher education in the area of recruitment, attrition and retention (eg, see EU, 2004; OECD, 2006; EC, 2007). Another important concern is the fact that a significant fraction of students starting a higher education programme in this area never complete their programme studies (OECD, 2008). In Sweden, only slightly more than half the students enrolling in an engineering master programme, with a nominal study time of 4,5 years, have completed their degree after seven years (SCB, 2008). Concerns in these areas have led to attributes related to recruitment, attrition and retention being increasingly used both to assess the quality of teaching and learning and to guide educational reform. In this scenario, enhanced understanding of the relations between the motives, academic achievement and retention become crucial to help prevent overly simplistic generalizations emerging from statistics in the area (Danaher, Bowser & Somasundaram, 2008). In countries like Sweden that have explicit widening of higher education access goals, because of the added complexity that this brings, such enhanced understanding becomes even more crucial.

Part of this understanding is underpinned by students' motives for enrolling in educational programmes and exploration of this has typically been undertaken using fixed-item questionnaires. We would argue that this approach generally takes for granted what motivational dynamics are critical for the design of the questionnaires. Hence we chose an alternative way of exploring and characterising student motives. We drew on Gee's (2005) concept of "Discourse Model" to identify the "implicit theories" students use to make sense of their motives for choosing to study in an engineering programme at a well regarded university in Sweden. Relations between these Discourse

Models and academic achievement and retention in the first year of the programme are given and then discussed in terms of proposed ways to enhance the teaching and learning in such programmes.

Rationale and theoretical framework

The social aspect of achievement and retention

Research on attrition and retention has shown that the social aspects of participation in higher education play an important role in the formation of students' academic trajectories. Here, the central theoretical modelling for student persistence in higher education has been done by Tinto (for example, 1975, 1987, 1993 -- *the Student Integration Model*) and by Bean (for example, 1982 -- *Student Attrition Model*). Although, at one level, significant differences between these models can be identified, the models share similar framing and thus can be seen to be largely complementary (Cabrera, Castenada, Nora & Hengstler, 1992). Our study draws upon some of this complementation; success in higher education is largely dependent on a positive interaction between the individual student and the discourse of the university and the educational programme.

Discourse models as an analytical tool

Discourse analysis is a widely used in qualitative educational research (for example, Wickman & Östman, 2002) and a wide variety of approaches to discourse analysis can be found in the literature. James Gee (2005) has developed a fruitful approach to such analysis and methodologically, Gee's work provides an integrated perspective, which we posit is particularly valuable for exploring the interaction between students and educational "Discourse". Gee's framework is built on a "family" of approaches, which have been drawn on vis-à-vis a seeking to illuminate the significance and implications of social, cultural, and political practices. Here, a notion of particular relevance to our study is Gee's notion of "Discourse Models". In Gee's modelling, Discourse Models are "simplified theories" that are used to make sense of the world and personal experiences in it (cf Gee 2005). Thus the identification of student Discourse Models provided us with a useful way to explore relations between students' identity, higher education experience, and academic achievement, in terms of students' perceptions of what is "valuable" and/or "normal".

Research questions

Drawing upon this theoretical background, two research questions regarding student motives, achievements and retention on a master programme in engineering physics were formulated for the study.

- Firstly, in terms of Discourse Models, what different motives for selecting a master programme of engineering physics can be identified?

- Secondly, do these Discourse Models relate to student academic achievement and programme retention, and if so how do they relate to each other?

Methods

The student population investigated in this study consisted of a cohort of first-year students on a master programme in engineering physics at a traditional research university in Sweden. 67 of the 89 students actively starting the programme choose to participate in the study. A combination of qualitative and quantitative approaches were used for this study: the identified Discourse Model characterizations are used to establish relations between academic achievement and programme-selection motivation.

Questionnaire study about motives

The data regarding student motives was collected using an *open-ended* questionnaire that was distributed to the students during their first week at university. The themes that made up the open-ended categories in the questionnaire were developed through earlier trials with groups of similar-profile students where broad questions were asked about student motives and expectations.

The descriptions given by students in the resultant open-ended questionnaire were analyzed using an iterative sorting procedure. This started with identification and coding of recurring threads found in the data and continued into a Discourse Model analysis. Four distinctly different Discourse Models describing student reasoning for enrolling to the programme were identified and categorized. These were, then, used to identify the dominant Discourse Model describing each student's motivational storyline as a basis for establishing relations between academic achievement and programme-selection motivation.

Study achievements

For the exploration of relations between motives and academic achievement, the course assessment results for the first year of study were obtained from the study records. The number of obtained "study-credits" and the student's choice of continued enrolment during the first year were taken to represent their academic achievement. This allowed a classification of students into three different categories. (1) *Successful:* where the student remained on the programme and attained results that met or exceeded official university modelling of a "successful" first year of study (≥35 ECTS). (2) *Unsuccessful*: where the student remained on the programme, but did not match the model of a "successful" first year of study (<35 ECTS). (3) *Leaver*: where the student left the programme during or after the first year of study. A distribution of "achievement categories" was then obtained for each student group as identified by a given Discourse Model.

Validity of results

The qualitative analysis process was continuously informed by discussion with colleagues. The resulting Discourse Models were presented to and discussed with groups of programme students, alumni and the programme advisory board. Their feedback provided a strong level of validity for the identified Discourse Models and greatly informed the discussions about the results. A second researcher was consulted to repeat and validate the identification of individual student's dominant Discourse Models.

Results

Motives for programme choice

Four different Discourse Models were identified in the discourse analysis and we labelled these as *Programme student, Engineer to be, Cosmic explorer* and *Convenience student*. A general storyline was constituted for each of these Discourse Models to capture the general features found in the student answers.

Engineer to be

The general storyline for the *Engineer to be* Discourse Model is that the programme is seen as a good preparation for a future career in the technical sector, probably as an engineer.

This view is in very good agreement with the intended goal of the programme to be a vocational programme for future engineers, and is generally shared by programme coordinators, study councillors and other academic staff. However, the quantitative analysis shows that only 12 out of 67 students could be characterized by this Discourse Model.

Cosmic explorer

Students characterized by the *Cosmic explorer* Discourse Model believe that the programme will help them to better understand the universe. In other words, these students were curiosity-driven. They all had had positive previous experiences of physics studies and envisaged a future career as a researcher. This Discourse Model characterized 11 of the 67 students.

Programme student

The *Programme student* Discourse Model encompasses a general and positive view of the programme. It is expected to give an enjoyable study experience and broad possibilities for the future. There will be no problem to get work after a completed programme, but the idea of what that job will entail is very vague. In this Discourse Model the excellent reputation of the programme, which carries a broad course base, was perceived to guarantee that "one can't really go wrong by choosing this programme, even if one doesn't have a clear plan for the future". The students characterized by this Discourse Model see the programme as being a part of a natural study progression. This was the most common of the four Discourse Models, characterizing 37 of the 67 students.

Convenience student

The students characterized by the *Convenience student* Discourse Model all see higher education as something one is expected to do, both by family and society in general. As these students had no clear idea on which programme to choose, they elected to enrol in a programme they perceived as convenient and not carrying too heavy a work load nor the content being particular difficult for them to master. This choice was usually based upon perceptions of physics and mathematics as easy subjects. After the programme these students future plan was to "relax, travel or just hang around". This Discourse Model characterized 7 of the 67 students.

Relations between motives, academic achievement and retention

The distributions of academic achievement categories within the student groups characterized by the four different Discourse Models are presented in Figure 1 below. As can be seen, there is a significant difference between the distributions.

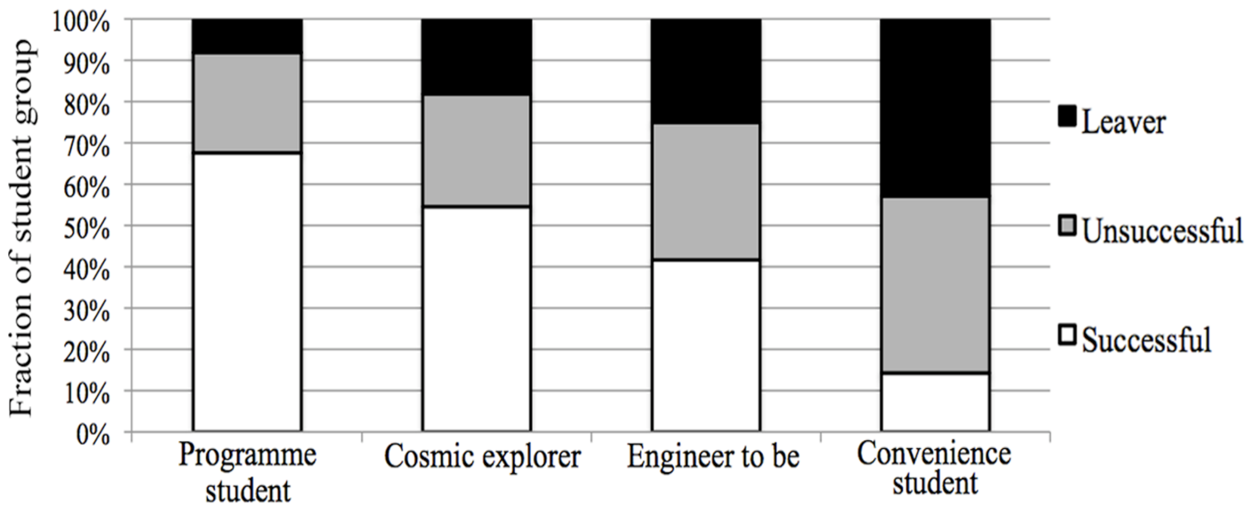


Figure 1. Distribution of academic achievement profiles.

The most successful group are the students characterized by the *Programme Student* Discourse Model. It is particularly interesting to note that a majority of the *Engineer to be* group have serious problems in terms of achieving academic success in their programme of study (even though these students are seen as most closely exemplify the target group of the programme by academic staff). That the *Convenience student* group, who thought that the programme would be "easy", had very big problems succeeding on this challenging programme, was, however, in complete agreement with our expectations.

Conclusions and Implications

We reflect on the results of this study in terms of a changing higher education landscape in a modern society, where an increasing fraction of the population is attending higher education and where some curriculums and associated teaching practices "mismatch" (McDermott 1993) this widening of access. An increasingly multi-faceted student population will bring with it a wider diversity of selection motivation. Such motives may, consequently, differ profoundly from what the programme and curriculum designers, and the teachers in the programme, arguably would be expecting. This is exemplified in our results by the Discourse Models of the *Programme students* and the *Convenience students*. The students characterized by these Discourse Models are primarily enrolling in higher education to become university students with no particular career goal in mind. Administrators and educators would see this is a short-term perspective, but it is well in line with some of the changing views on education found in contemporary youth culture (Miles, 2000).

Hence, an increasingly multi-faceted student population presents challenges with regard to a programme's design, realization, knowledge of students as learners, and approaches to teaching. For example, the programme in our study is explicitly designed and conceptualized as a vocational programme aimed at producing high quality engineers. Thus it is envisaged to attract the vocationally minded student, yet relatively few students in the studied cohort of first-year students entered the programme with a motive that fitted this vocation identity. Furthermore, a large proportion of those that did not have this motivation actually have a higher probability to be successful: the *programme student* and the *cosmic explorer* student both had higher successful profiles than the *engineer to be* students. A very unexpected and disturbing outcome! This trend is something that programme and curriculum developers, teachers, and researchers need to collaboratively investigate and attempt to address in order to craft an educational environment that offers a much better possibility of academic success for all students, and better possibilities for an associated development of their Discourse Model.

References

Bean, J.P. (1982). Conceptual models of student attrition: How theory can help institutional researchers. I E. T. Pascarella (Ed.) Studying Student Attrition, San Fransisco:Jossey-Bass.
Cabrera, A. F., Castaneda, M. B., Nora, A. & Hengstler, D. (1992). The Convergence between Two Theories of College Persistence. *Journal of Higher Educ.,63*(2), 143-164.
Danaher, P.A., Bowser, D. & Somasundaram, J. (2008) The student departure puzzle: do some faculties and programs have answers?, *Higher Education Research & Development* 27(3), 271-280
EC (2007), Science Education NOW: A Renewed Pedagogy for the Future of Europe, EUR22845, European Commission, Luxembourg
EU (2004), Europe needs more scientists!, European Commission, Brussels
Gee, J.P. (2005) An introduction to discourse analysis. Theory and Method. Second edition. London: Routledge.
McDermott, L. (1993) Guest Comment: How we teach and how students learn—A mismatch? *American Journal of Physics* 61, 295-298
Miles, S. (2000) Youth lifestyles in a changing world. Buckingham: Open University Press.
OECD (2008). Education at a glance - 2008, OECD
OECD (2006). Evolution of Student Interest in Science and Technology Studies - Policy Report: Global Science Forum, OECD
SCB (2008), Utbildningsstatistisk årsbok 2008, Statistiska Centralbyrån, Stockholm
Tinto, V. (1975) Dropout from higher education; a theoretical synthesis of recent research, *Review of Educational*

Research, 45(1), 89-125
Tinto, V. (1987). Leaving College: Rethinking the causes and cures of student attrition. Chicago: University of Chicago Press.
Tinto, V. (1993) Leaving College: Rethinking the Causes and Cures of Student Attrition. Chicago:University of Chicago Press.
Wickman, P.-O. & Östman, L. (2002) Learning as Discourse Change: A Sociocultural Mechanism. *Science Education* 86(5), 601-623.

EXAMINING CONCEPTUAL CHANGE ON THE GREENHOUSE EFFECT AND GLOBAL WARMING USING COGNITIVE CONFLICT STRATEGY IN A CONSTRUCTIVIST LEARNING ENVIRONMENT

Sibylle Reinfried
University of Teacher Education of Central Switzerland (PHZ) Lucerne

Abstract

Many people attribute the greenhouse effect and the resulting global warming to a hole in the atmosphere. This "hole paradigm" and other naïve mental models about the causes of the greenhouse effect are extremely resistant to educational efforts and difficult to change. In order to be able to support political measures to reduce the greenhouse gas CO_2, citizens must grasp the fact that anthropogenic CO_2 is the major cause for global warming. A longitudinal survey using a pretest-posttest design with intervention and a delayed posttest after two months was used to analyze the efficacy of a conventional teaching method in comparison with a constructivist learning environment. The constructivist setting geared to the implementation of a long-term conceptual change from naïve mental models to appropriate scientific concepts by employing a cognitive conflict strategy. This paper reports about the pilot study in which a total of 41 14-year-old students (two comparative classes of 8 graders; 21 girl, 20 boys) participated. The results of the statistical analysis show that the intervention caused a significant increase in knowledge and a better understanding of the principle of the greenhouse effect in the experimental group in the posttest and the delayed posttest. This suggests that learning environments that provide immediately persuasive and plausible experiences for learners are better suited to initiate conceptual change than conventional teaching methods in the lower secondary science classroom.

Background

With a view to the very real threat of global warming (IPCC, 2001, 2007), it is imperative that young people understand the origins and impacts of the greenhouse effect. However, environmental concern of a general nature appears not to motivate people to support programs designed to control global warming. It is the correct understanding of its causes that is seen as the key trigger of behavioral change to address global warming (Bord et al., 2000; O'Connor et al., 1999; O'Connor et al., 2002). Already in the early 1990s, Aeschbacher (1992) and Boyes and Stanisstreet (1993) found in their investigation of children's and adults' preconceptions on the greenhouse effect that many individuals were linking the phenomena of ozone layer depletion and global warming. Particularly widespread are the 'greenhouse model' and the 'ozone hole model'. The greenhouse model is a scientific-like mental model, according to which the rays of the sun reach the Earth unimpeded through the atmosphere, but the reflected radiation can no longer leave the atmosphere, because exhaust gases lead to the formation of a layer in the upper atmosphere. In the ozone hole model a hole in the atmosphere allows more sun rays to enter, which are reflected by the surface of the Earth and then can no longer find the hole (meaning, the exit), which makes the atmosphere become warmer (Schuler, 2005; Reinfried et al., 2008). Attempts to impart a correct perception of the GHE by means of linguistic or symbolic instruction met with little success (Gauthier et al., 2006). According to several longitudinal studies (Aeschbacher et al., 2001; Anil, 2003; Eisele, 2003), this holds especially true for the retention rate of the cognitive gains acquired in the learning process.

Research Questions and Problem

In the present pilot study the understanding of the greenhouse effect is regarded as resulting from a very stable framework theory, which is formed at an early age. It is not accessible to conscious awareness and hypothesis testing and is deeply entrenched in an individual's everyday life (Vodniadou & Brewer, 1992; Vosniadou, 1994). The "hole paradigm" to explain the greenhouse effect is constructed on the basis of the everyday observation that "something can come in through a (ozone) hole" in combination with the naïve concept "sun rays produce heat". Laypeople automatically equate transparency of gas with transmissivity of radiation. On the other hand, the phenomenon of "heat absorption in CO_2" is non-existent in our everyday experience. In addition, analogies and metaphors play an important role when individuals construct their mental models about the greenhouse effect. Such analogies are the glass hothouse and the way atmospheric processes are visualized graphically in books, journals or on the Internet (Reinfried et al., in print). The notion that radiation is captured under the (glass) layer of a hothouse and the lack of experience with the heat absorption of gases lead to the construction of the greenhouse model.

Re-learning in the sense of conceptual change can happen when a new concept induces a cognitive conflict in the learner (Posner et al., 1982). For a learner to understand that the "hole paradigm" or the "greenhouse paradigm" cannot be upheld, the following physical phenomena and their interrelatedness must be understood:

1) Solar radiation is for the most part absorbed and not reflected by the Earth's surface. The Earth's surface then reradiates the "absorbed" energy in a converted form upwards – not in form of light but in form of heat radiation.

2) CO_2 molecules in the atmosphere absorb the radiation emitted by the Earth and reradiate it in all directions. Thus, they slow down the emission of heat energy into space. Solar radiation, however, passes the CO_2 molecules unhindered.

On the basis of these considerations, a constructivist learning environment consisting of worksheets and a model experiment was developed (Reinfried et al., 2008). It

a) takes into consideration the above-mentioned presuppositions of the origins of the greenhouse effect.

b) combines the preconditions for a conceptual change according to the cognitive conflict approach enabling the reorganization of the underlying assumptions (Posner et al., 1982).

c) incorporates measures for the optimization of text and images based on the cognitive theory of multimedia learning (Kintsch, 1998; Mayer, 2001; Pavio, 1986).

In this pilot study (and the following main study with 335 students) the following research questions were investigated: How big is the knowledge gain in relation to the targeted conceptual change under various conditions - the experimental group using study material, which was optimized according to the findings of the psychology of learning and the control group using conventional material?

Methods

A pretest(t1)-posttest(t2) design with intervention and delayed posttest (t3) after eight weeks was used. The sample consisted of two groups of 20 resp. 21 eighth graders from a big town in German-speaking Switzerland (a total of 41 14 year olds, 21 females, 20 males), students who have not yet had instruction in physics or chemistry. A questionnaire was administered in the pre- and posttest to measure the increase of students' level of knowledge. The questionnaire consisted of 32 knowledge questions answered on a five-point Likert response scale. In addition, students were asked to draw sketches and describe in words their conception of the greenhouse effect. The sketches and descriptions were then codified and thus quantified. Students' understanding before and after the intervention was calculated by adding up points given for the correct answers to the knowledge questions on the questionnaire, as well as for the sketches and their captions. In the intervention, which lasted 60 minutes, students worked in a learning environment that consisted of work sheets and an experiment. The design of the intervention used the same teaching strategies such as direct instruction, single work and work in pairs, hypothesis testing, an experiment and a group discussion in both groups. The work sheets contained the same information for both groups. While instructions for the experimental group were optimized according to instructional psychology, those for the control group consisted of pictures and texts taken from conventional textbooks for the 8th grade. The experiment in the experimental group showed heat absorption by CO_2 only in a qualitative, but easy-to-follow way in order to trigger a cognitive conflict. In this experiment CO_2 was filled into a transparent container standing in the path of heat rays - emitted from a globe's surface - and an infrared radiation meter. Meter reading indicated that the invisible CO_2 gas obstructed the invisible heat rays emanating from the "Earth's" surface. The experiment for the control group concerned the absorption of heat rays by gases according to Tyndall (Parchmann et al. 1995; Parchmann & Jansen 1996). It demonstrated the same effects quantitatively according to the methodological criteria of exact experimental settings in science instruction adapted to the lower secondary level.

Results and Discussion

Reliability analyses of the questionnaire items revealed low to medium scores for the knowledge tests (Cronbach's α_{t1}= .59; α_{t2} = .80, α_{t3} = .85). The relatively low score in the pre-test was a result of students' little pre-instructional knowledge. The drawings and descriptions of the pre-test evinced that the "hole-paradigm" (40%) and the "greenhouse-paradigm" (31%) were widespread in both groups. After the intervention, a knowledge gain could be observed in both groups. In the experimental group, however, this knowledge gain was significantly larger in the posttest immediately after the intervention and in the delayed posttest after eight weeks (M_{t3EG}= .73, SD = .23; M_{t3CG}= .54, SD = .19; $F(1,34) = 10.61$, $p > .01$, $Eta^2 = 0.24$) (Fig.1).

Students in the experimental group demonstrated either a partial or far-reaching conceptual change relating to the significance of solar radiation conversion for the greenhouse effect. In the delayed posttest still 65% of the students answered the questions relating to the "hole-paradigm" correctly. 55% of the students exposed correct knowledge concerning the significance of CO_2 as a greenhouse gas. The knowledge gain in the experimental group was the result of a fundamental conceptual change displayed in students' drawings and texts. 50% of the students in the experimental group, in contrast to only 20% in the control group, had grasped the major physical concepts causing the greenhouse effect. The main perplexing factor for the learners was that the invisible CO_2 was blocking, to a large extent, the heat rays emanating from the Earth model. Students could experience the fact that while CO_2 allows light to pass through, it absorbs heat rays. Re-learning has occurred in spite of the "resistance" of the seemingly plausible but scientifically untenable "hole-paradigm". The control group, too, has experienced some knowledge gain. However, materials taken from conventional textbooks explain the processes in the atmosphere in a manner that cannot readily be understood and is not sufficiently plausible. The experiment that showed the heat ray absorption quantitatively required learners to produce a considerable motivational and cognitive personal effort in order to transfer the knowledge gained from the experiment to the realities of the Earth's atmosphere. It may therefore be assumed that this type of instructional setting may have only induced a superficial knowledge rather than a permanent conceptual change.

Conclusions and Implications

These results imply that, at the lower secondary level, instructional constructivist settings that are simple, attractive, tangible, readily understandable, and convincing are better suited to laying the foundation for understanding complex and abstract environmental problems. In contrast, conventional teaching material and physical experimental settings, while allowing exact measurement, seem to require considerable effort by the learner to transfer conceptual understanding to the real problem. When developing learning environments intended to initiate conceptual change, one should therefore have the courage to harness the relevant special branch of science (in this case: physics) for the good of the targeted phenomenon (here: the greenhouse effect) and to subordinate exact experimentation to learning experience.

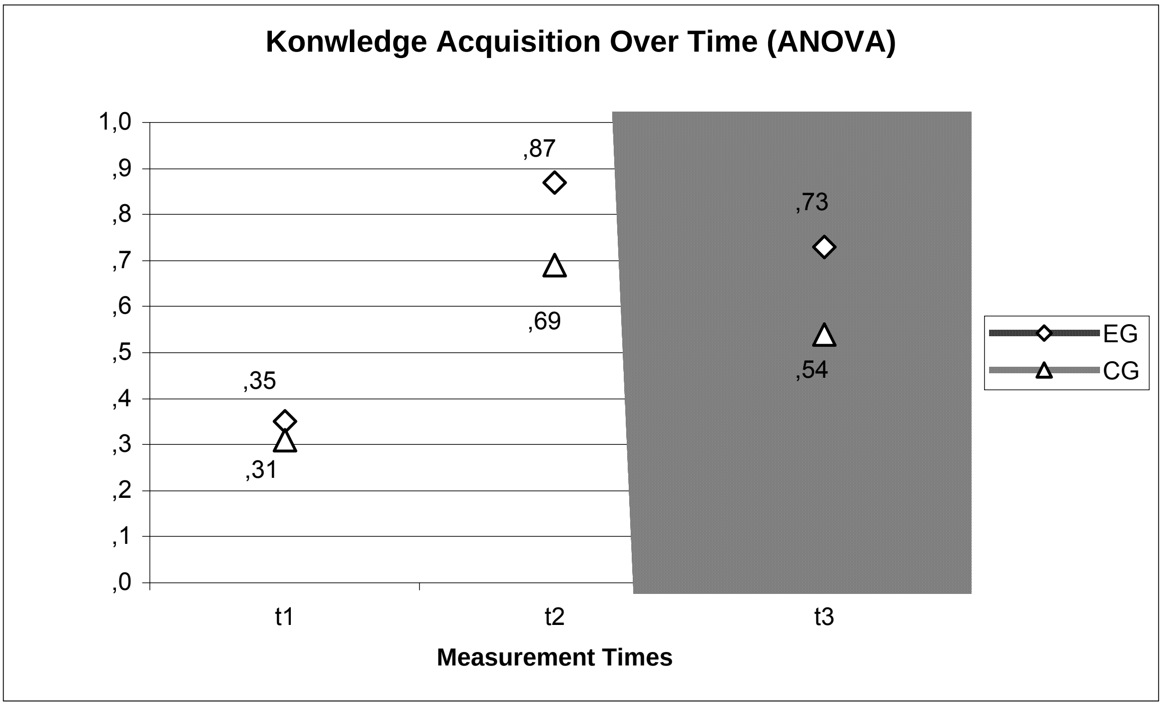


Figure 1. Knowledge acquisition of the two groups over time (EG = experimental group; n=19, CG = control group; n=17)

According to Bord et al. (2000), O'Connor et al. (1999) and O'Connor et al. (2002) the knowledge and correct understanding of what causes and what does not cause the greenhouse effect is the strongest single predictor, both of behavioral intention of seizing voluntary action and of supporting political measures to greenhouse gas reduction. People who imagine that a hole in the atmosphere or exhaust gases are the reasons for global warming want to do something to combat the destruction of the ozone layer, e.g., no further use of spray cans - although the use of long-lived fluorochlorinated hydrocarbons (FCHCs) has been banned in the EU since 1996 and in Switzerland since 2004 - and about air pollution, meaning smog, smoke, particulate matter and odorous substances (Kempton, 1993). Their conduct as consumers and voters thus remains ineffective with regard to the actual objective of reducing CO_2 emissions. Only the knowledge and understanding that

1. it is not a hole in the atmosphere but the impact of the CO_2 'ray catcher' that is the cause of the greenhouse effect, and

2. the anthropogenic greenhouse impact on the increase in the CO_2 content of the atmosphere is based primarily on coal and petroleum, caused by the burning of fossil fuels,

make it possible for people to perceive the interdependencies between the greenhouse effect and global warming and helps to identify CO_2-reducing measures. They are the foundations upon which the insight and readiness can be developed to reduce one's own CO_2 emissions or to support political measures for the reduction of CO_2.

References

Aeschbacher, U. (1992). Meinungen, Wissen und Verstehen von Lehrerstudentinnen und –studenten in Sachen „Treibhauseffekt“. *Bildungsforschung und Bildungspraxis* 14(2): 149-161.

Aeschbacher, U., Calò, C., Wehrli R. (2001). „Die Ursache des Treibhauseffektes ist ein Loch in der Atmosphäre“: Naives Denken wider besseres Wissen. *Zeitschrift für Entwicklungspsychologie* 33(4): 230-241.

Anil, D.(2003). *Dauerhafte Korrektur mentaler Fehlkonzeptionen am Beispiel des Treibhauseffekts. Textverstehen und Textverständlichkeit.* Unpublished master thesis. (Basel: Institut für Psychologie, University of Basel).

Bord, J. R., O'Connor, R.E., Fisher, A. (2000). In what sense does the public need to understand global climate change? *Public Understanding of Science* 9: 205-218.

Boyes, E. & Stanisstreet, M. (1993). The "greenhouse effect": children's perception of causes, consequences, and cures. International Journal of Science Education 15(5): 531-552.

Eisele, C. (2003). *Dauerhafte Korrektur mentaler Fehlkonzeptionen am Beispiel des Treibhauseffekts. Analogie und Abbildung. Unpublished Master thesis* (Basel: Institut für Psychologie, University of Basel).

Gautier, C. et al. (2006). Misconceptions about the Greenhouse Effect. *Journal of Geoscience Education* 54(4): 386-395.

I PCC (2001). *Climate Change 2001: The Scientific Basis. Contribution of the Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change.* (Cambridge University Press, www.ipcc.ch (last accessed on 15th August 2009).

IPCC. (2007). Summary for policymakers. In S. Solomon, D. Qin, M. Manning, Z. Chen, Z., M. Marquis, K.B. Averyt, M. Tignor & H.L. Miller (Ed.), *Climate Change 2007: The physical science basis. Contribution of working group I to the fourth assessment Report of the Intergovernmental Panel on Climate Change.* (Cambridge, United Kingdom, and New York: Cambridge University Press).

Kempton, W. (1993). Will public environmental concern lead to action on global warming? *Annual Review of Energy and the Environment,* 18, 217-245.

Kintsch, W. (1998). *Comprehension, a pardigm for cognition.* (Cambridge: Cambridge University Press).

Mayer, R. E. (2001). *Multimedia Learning.* (Cambridge: Cambridge University Press).

O'Connor, R. E., Bord, R. J. & Fisher, A. (1999). Risk perceptions, general environmental beliefs, and willingness to adress climate change. *Risk Analysis,* 19(3): 461-471.

O'Connor, R. E., Bord, R. J., Yarnal, B., Wiefek, N. (2002): Who wants to reduce greenhouse gas emissions? *Social Science Quarterly*, 83(1): 1-17.

Parchmann, I., Kaminski, B. & Jansen W. (1995). Die Wärmeabsorption von Gasen. *CHEMKON,* 2 (1), 17-25.

Parchmann, I. & Jansen W. (1996). Der „Treibhauseffekt" als Folge der Wärmeabsorption von Gasen. *CHEMKON* 3(1): 6-11.

Pavio, A. (1986). *Mental representations. A dual coding approach.* (Oxford: Oxford University Press.)

Posner, G., Strike, K., Hewson, P. & Gertzog, W. (1982). Conceptual change and science teaching. *European Journal of Science Education*, 4 (3): 231-240.

Reinfried, S., Schuler, S., Aeschbacher, U. & Huber, E. (2008): Der Treibhauseffekt – Folge eines Lochs in der Atmosphäre? Wie Schüler sich ihre Alltagsvorstellungen bewusst machen und sie verändern können. *geographie heute,* 265/266, 24-33.

Reinfried, S., Rottermann, B., Aeschbacher, U. & Huber, E. (in print). Alltagsvorstellungen über den Treibhauseffekt verändern und so die globale Klimaänderung verstehen. *Geographie und ihre Didaktik 3.*

Schuler, S. (2005). Umweltwissen als subjektive Theorie: Eine Untersuchung von Schülervorstellungen zum globalen Klimawandel. In: M. Schrenk & W. Holl-Giese (Hrsg.), *Bildung für eine nachhaltige Entwicklung - Ergebnisse empirischer Untersuchungen* (S. 97-112). Hamburg: Kovac.

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. Learning and Instruction 4: 45-69.

Vosniadou, S. & Brewer, W. F. (1992). Mental models of the earth. Cognitive Psychology 24: 535-585.

MOTIVATION TO LEARN SCIENCE AND COGNITIVE STYLE

Sibylle Reinfried
Universität Zürich

Abstract

Whilst gender is considered to be one of the most significant factors influencing attitudes towards science, previous research seems to suggest that at least in non science classes there is no correlation between gender and motivation to learn science. The present study investigates into a mixed group of science and non science students of upper secondary level. The data show that there is in fact no correlation between gender and motivation to learn science, but there is a highly significant positive correlation between the students' cognitive style and their motivation towards science learning. A systemizing cognitive style supports a high motivation to learn science, and vice versa. At the same time – as also known from previous research - male students show a more systemizing cognitive style whilst female students present a more empathizing one. These results suggest that the cognitive style is a basic variable of motivation to learn science, while the impact of gender is of second order.

Introduction

Gender is one of the most significant factors influencing attitudes towards science (Osborne 2003). Numerous research studies show that boys have consistently more positive attitude to school science than girls, and that this effect is stronger for physics than for biology, and highest for "general science" which does not distinguish between different science subjects (p. 1062). However what concerns the motivation to learn science, the situation seems to be more unclear. Whilst the research literature on science majors suggests that gender has an important impact on the learning of science in science courses (Green & DeBacker, 2004), much less is known about the role of gender in the science learning of non science majors. Glynn and colleagues (2007) could in fact not find any correlation between gender and motivation to learn science in this group. They interpreted this unexpected result in terms of the special conditions in science courses for non science students: "…In science courses for non science majors there is a relatively 'level playing field' that supports the women's motivation as well as the men's. We believe our finding generalizes across science courses for non science majors". (p. 1101).

The here presented research is committed to a strand of research coming from cognitive science, which suggests another explanation. Billington and colleagues (2007) investigated into students in physical sciences and humanities. They found that sex was indeed a predictor for the entry either into physical sciences resp. humanities. However the so called cognitive style, characterized by systemizing resp. empathizing activities, was much better as a predictor than gender.

These results suggest that the cognitive style as defined in the research of Billington and colleagues could also be a better predictor of the motivation to learn science than gender is.

The research questions of our study therefore were twofold:

1) Can we find an influence of gender on the motivation to learn science, and this also in a mixed group of science and non science students?

2) Can we find an impact of the cognitive style on this motivation in the same group of students?

Our research hypothesis was that there is an influence of one's cognitive style on the motivation to learn science, but that in a mixed group of students there is no influence of gender on this motivation. By the verification of this hypothesis our study tries to add one more piece to the puzzle. If the cognitive style influences the motivation towards learning science, and, as Glynn and colleagues show, motivation influences achievement in science subjects, then this could explain why a systematizing brain type predestinates for physical science degree studies. Apart from that, this type of results would also confirm and expand the results of Glynn and colleagues which did not find a correlation of motivation and gender in non science major classes.

Rationale

Motivation

In a review of the major literature about attitudes to science and its implications Osborne et al. (2003) report that studies measuring the attitude towards science have incorporated motivation towards science as one component of others including the perception of the science teacher, anxiety towards science, the value of science, self-esteem at science, enjoyment of science, attitudes of friends and peers towards science, attitude of parents towards science, the nature of the classroom environment, achievement in science, and fear of failure on course.

In the here presented study, to maintain the comparability of the results of this study with those of Glynn and colleagues, their theoretical framework of motivation was used. Thus motivation is defined as "…the internal state that arouses, directs, and sustains students' behaviour towards achieving certain goals. In studying the motivation to learn science, researchers attempt to explain why students strive for particular goals, how intensively they strive, how long they strive, and what feelings and emotions characterize them in this process." (S. 1090). Based on research within the social-cognitive motivational framework (Bandura, 2001), the authors identify five important motivational constructs that include intrinsic and extrinsic motivation, namely goal orientation, self-determination, self-efficacy, and assessment anxiety (Glynn & Koballa, 2006). The so–called Science Motivation Questionnaire (SMQ) (see chapter 4) reflects these five motivational constructs. Motivation towards science, as defined by these five constructs, overlaps with some of the components of the attitude towards science as determined in Osborne et al. (2003). However, the in-principle idea, namely that motivation towards science is an important component of the attitude towards science, which is a more complex, overarching concept, remains in place.

Cognitive style

The approach of cognitive styles used by Billington and colleagues is based on a recent theoretical account of cognitive style differences of Baron-Cohen and colleagues (Baron-Cohen, Knickmeyer, & Belmonte, 2005). It proposes two core psychological dimensions, or cognitive styles: empathizing (E) and systemizing (S) (Billington et al., 2007).

Systemizing is defined as a drive and ability to analyse the rules underlying a system, in order to predict its behaviour. A system in this context is understood as an object showing a tripartite structure: It can always be analysed in terms of so-called input – operation –output patterns, where inputs are initial states of the system, outputs as subsequent states of the system, and operations as actions that transform input states into output states. Defined in this general way, systems can be found in many different domains: technical (e.g. machines and tools); natural (e.g. weather system); abstract (e.g. mathematics); social (e.g. political system); spatial (e.g. map reading); and organisable (e.g. a taxonomy). A systemizing view on objects of interest is able to understand these objects in terms of a system, which needs an ability to identify local details and their interaction and to abstract from Gestalt perceptional distracters, also known as a "field independent" cognitive style (Witkin, Dyk, Faterson, Goodenough, & Karp, 1962).

Empathizing is defined as a drive to identify another's person mental states and to respond to these with one of a range of appropriate emotions. Empathizing has thus both a cognitive and an affective component (Baron-Cohen & Wheelwright, 2004; Davis, 1980). The cognitive component involves understanding another's person thoughts and feelings and is also referred to as using a theory of mind (Wellman, 1990). The affective component of empathizing involves an emotional response that arises as a result of the comprehension of another individuals emotional state (Eisenberg, 2002).

Every human being is considered to dispose of both of these cognitive styles, empathizing and systemizing, but normally on a different level. Some individuals are rather systemisers (S>E) whilst others have a dominant empathizing cognitive style (E>S). Others show a balanced type (E=S) of cognitive styles. The relation of E and S is called the brain type of the individual. The whole concept is called the E-S model.

In order to work with the E-S model, two self-reporting questionnaires have been developed and tested (see chapter 4). The two questionnaires exist in different versions, but each of these calculate a systemizing quotient (SQ) and an empathizing quotient (EQ) providing a measure of the individual's capacity to use the two cognitive styles. The variable representing the brain type is essentially calculated as the normalized difference of EQ and SQ.

One of the important research results based on these questionnaires is that females on average have a stronger drive to empathize (E>S), whilst males on average have a stronger drive to systemize (Baron-Cohen, 2002). This claim only applies on average; thus there will always be individuals who are atypical for their sex. However, and this is the important point in the context of our study, the E-S theory also argues that, irrespective of their sex, if an individual's systemizing is at a higher level that their empathizing (S>E), it is this profile that leads them into disciplines that require an analytical style to deal with rule-based phenomena (Billington et al., 2007). If this were true, then it could be that it is basically not gender that mostly influences the attitude towards science subjects, but the brain type which only on average coincides with the individual's sex.

It is in this theoretical framework that two recent studies (Billington et al., 2007; Wheelwright et al., 2006a) indeed showed, that physical science degree students scored significantly lower on the EQ and significantly higher on the SQ and suggested, that the academic subject one ends up studying may be better predicted by one's cognitive style than by one's sex.

Methods

At a science learning centre of a university in Switzerland we studied 77 upper secondary students (43 women and 33 men) being from 15 to 20 years old. Their mean age was m_{age}= 17.56 (SD=1.11).

The science learning centre offers one day courses to upper secondary classes of the region about life science topics. The teachers enrol the classes for these events because of different reasons, mainly as an excursion matching the curriculum, as a class event, or because they do not dispose of the experimental facilities at their school.

In Switzerland, in upper secondary school students cannot yet be classified as science or non science students. Every student has to take part in any subject of science and non science disciplines. However these students decide for so-called specializing issues, where they enjoy a higher education. In this sense, our sample included a broad spectrum of different students. 27 of our students were specialized in Music Studies (35.1%), 23 in Biology/Chemistry (29.9%), 17 in Modern Languages (22.1), 5 in Mathematics/Physics (6.5%), 4 in Ancient Languages (5.2%)) and 1 in Pedagogy/Psychology/Philosophy (1.3%). If we summarized students specialized in Mathematics/Physics resp. in Biology/Chemistry under the label of "science", and the other ones under the label of "non science", then we disposed of 28 science students (36.4%) and of 49 non science students (63.3%).

Procedure

The classes were visited during their one-day course at the learning centre. They were informed about the study and they consented to participate. Every student filled in one combined questionnaire and received his personal results by e-mail if he applied for.

The questionnaire

Part A, cognitive style

In part A of our questionnaire, we used the German version of the SQ and the EQ questionnaire by Baron-Cohen (Baron-Cohen, 2004). A pre-test showed that some of the questions had to be slightly modified to be useable for our students ("car" for example was replaced by "motorbike"). Both the SQ and the EQ questionnaire are 60-item, forced choice format, containing 40 cognitive style items and 20 control items. The SQ asks questions such as "I like music shops because they are clearly organized" and "When I learn a language I become intrigued by grammatical rules". Similarly, the EQ asks items such as "I am good at predicting what someone will do" to measure cognitive empathy or "I usually stay emotionally detached while watching a film" to measure the affective component of empathy.

On both the EQ and the SQ, participants are asked to respond "definitely agree, "slightly agree", "slightly disagree" or "definitely disagree", and approximately half the items are reverse scored to avoid response bias. Scores on both the SQ and the EQ range from 0 to 80.

An EQ from 0-32 is considered as low, 33-52 as average range (most women score about 47 and most men score about 42), 53-63 is above average, 64-80 is very high.

A SQ of 0-19 is considered as low, 20-39 as average (most women score about 24 and most men score about 30), 40-50 as above average, 51-80 as very high.

A "Brain Quotient" B was calculated for each participant following a method reported in Wheelwright and colleagues (2006a). To this end, EQ and SQ were standardized to E=(EQ-<EQ>)/80 and S=(SQ-<SQ>)/80, where <EQ>=44.3 and <SQ>=44.3 are the population means found in literature (Baron-Cohen, Richler, Bisarya, Gurunathan, & Wheelwright, 2003; Wheelwright et al., 2006b). The division by 80 reflects the maximal score of EQ and SQ respectively. The "Brain Quotient" B then represents a coordinate transformation defined by:
B=(S-E)/2 and
C=(S+E)/2

B essentially calculates the difference between E and S. If it is negative then (E>S), and vice versa. A brain type can be defined by classifying the Brain Quotient as follows: Those scoring in the lowest and highest 35% of the population are classed as Type E and Type S, respectively. The remaining individuals, between the 35^{th} and the 65^{th} percentile, defined Type B (balanced type).

Part B, motivation towards science

In Part B of the questionnaire, we asked students to respond to the 30 items of the Science Motivation Questionnaire (SMQ; (Glynn & Koballa, 2006)) The items were translated into German and also tested in a pre-test. The SMQ items (see Table 1) were developed based on the motivation concepts described earlier in this article. The SMQ items ask students to report on intrinsically motivated science learning (items 1, 16, 22, 27, and 30), extrinsically motivated science learning (items 3, 7, 10, 15, and 17), relevance of learning science to personal goals (items 2, 11, 19, 23, and 25), responsibility (self determination) for learning science (items 5, 8, 9, 20, and 26), confidence (self-efficacy) in learning science (items 12, 21, 24, 28, and 29), and anxiety about science assessment (items 4, 6, 13, 14, and 18). Typical items for this questionnaire are "I enjoy learning science" (item 1) or "Earning a good science grade is important to me" (item 7) or "I am confident I will do well on the science labs and projects" (item 21). Students respond to each of the 30 randomly ordered items on a 5-point Likerttype scale ranging from 1 (never) to 5 (always). The anxiety about science assessment items are reverse scored when added to the total, so a higher score on this component means less anxiety. The SMQ maximum total score is 150 and the minimum is 30. A score in the range of 30–59 is relatively low, 60–89 is moderate, 90–119 is high, and 120–150 is very high (Glynn & Koballa, 2006).

Results

We computed statistics results by means of the Statistical Program for the Social Sciences, version 15.0 (SPSS). Because we translated the questionnaires and (slightly) adapted them to adolescents, the testing of the reliability (internal consistency) of the used questionnaires was essential. Cronbach coefficient alpha were 0.849 for SMQ (30 Items). 0.877 for SQ (40 items), and 0.899 for EQ (40 items) indicating that 85%, 87%, and 90% respectively of the variance of the total scores on this questionnaires can be attributed to systematic variance. This means that the questionnaires have preserved their (already proofed) high internal consistency in the new context.

The impact of gender

On average, our students show a relatively high motivation for learning science (M=99.56, SD=17.40). The SMQ of the male students is higher (SMQ_M=104.03, STD=14.81) than the SMQ of the female students (SMQ_F=96.23, STD=18.56), but still high for both groups. The gender difference is not significant. The same holds, if the analysis is restricted to the science students. The EQ for both groups is in the population average (M_{EQ}=40.15, SD=11.36), but the female students (EQ_F = 43.81, STD=10.14) have a higher EQ than the male students (EQ_M=35.22, STD=11.17). This difference is highly significant (p<.001) The SQ for both groups is also in the population average (M_{SQ}=25.81, SD=11.56), but the SQ of the male students (SQ_M=30.41, STD=13.39) is higher than the SQ of the female students (SQ_F=22.40, STD=8.67). This difference is highly significant (p<.001), too.

The average brain quotient of our students is balanced (M=0.02, SD=0.12). However the braintype of the female students on average is negative, i.e. on the empathizing side (BQ_F= -.02, STD=.08), while the braintype of the male students on average is positive, i.e. on the systemizing side (BQ_M=0.08, STD=.13). This difference is also highly significant (p<.001).

The impact of the braintype

There is a medium positive Pearson correlation between SMQ and BQ (r=.315). This correlation is highly significant (p<.005 2-tailed). If for example the students are classified into three groups as described in the methods part - the systemizing group (S), the balanced type group (B), and the empathizing group (E) - then the SMQ of the systemising students (SMQ_{syst}=91.50, STD=17.30) is significantly higher than the SMQ of the empathizing students (SMQ_{emp}=103.25, STD=15.18), (p<.01) (Table 1). The bar charts (Fig. 1) reveal an interesting distribution, namely whilst more empathizers show low motivation to learn science, much more systemisers show a high SMQ.

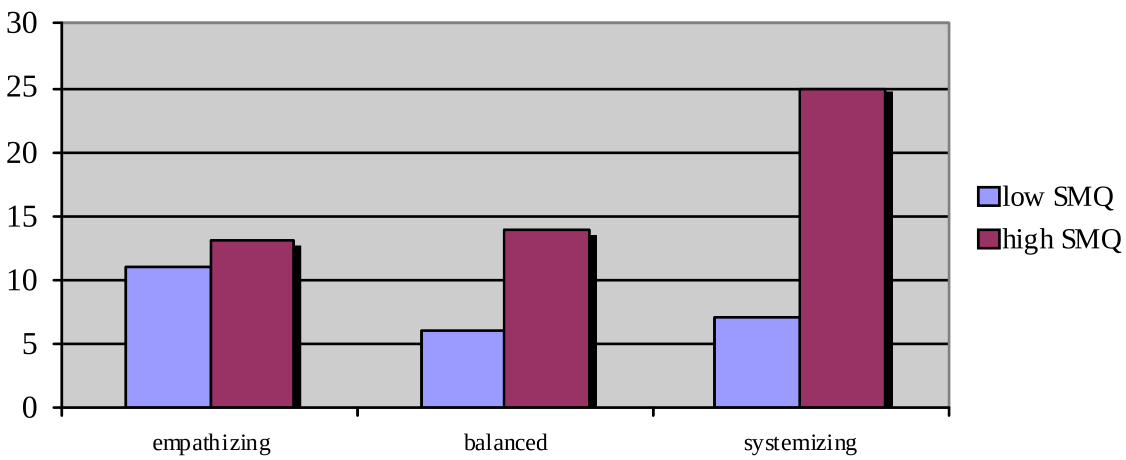


Figure 1. Bar Chart Braintype vs. Low/high Motivation

Additionally, the cross tabulation (Table 2) between science and non science majors and the braintype shows that there is also a higher percentage of systemisers in the science group (S_{Sci}=53.6%) than in the non science group(S_{Nonsci}= 36.2%), and a much higher percentage of empathizers in the non science group (E_{Nonsci}=44.7%) than in the science group (E_{Sci}=10.7%) (Figure 2). The Pearson Chi-Square test shows that this difference is highly significant (p<.005).

Table 1. Braintype * Low/high Motivation Cross tabulation

		Low/high Motivation		Total
		low	high	
Braintype	empathizing	11	13	24
	balanced	6	14	20
	systemizing	7	25	32
Total		24	52	76

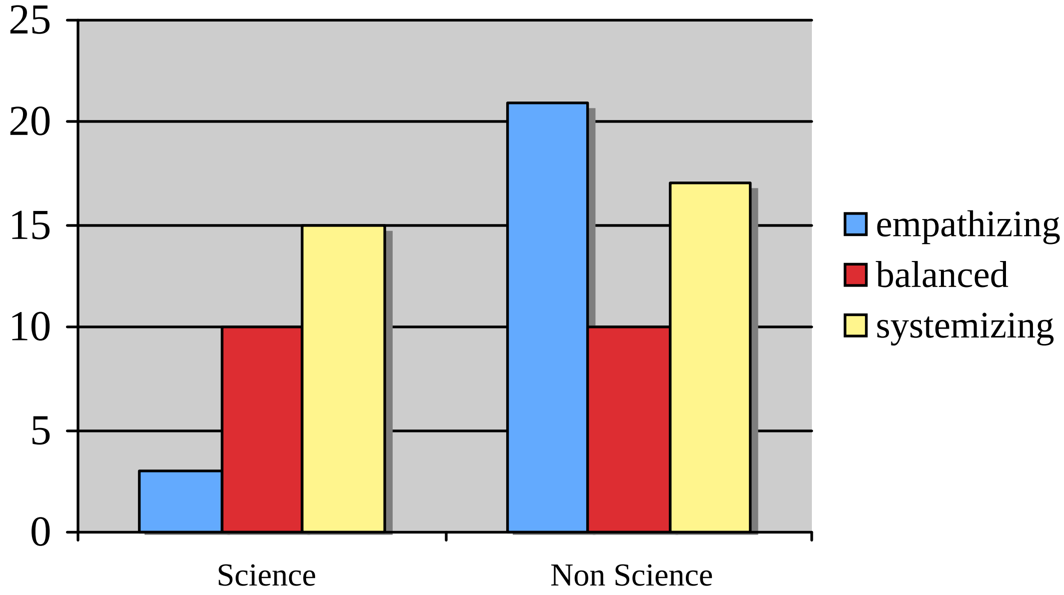


Figure 2. Bar Chart Science/Non Science vs. Braintype

The braintype is in fact a predictor of being a science or a non science student, while sex is not. This is shown by the data of a binary logistic regression. A model including these two variables is significant (Chi-square=6.53, p<.05), and overall 60% of the predictions of this model are accurate. The braintype is a predictor (W=5.17, df=1, p<.05), sex is not.

Table 2. Science/Nonscience vs. Braintype Cross tabulation

	Braintype			Total
	empathizing	balanced	systemizing	
Science	3	10	15	28
Non Science	21	10	17	48
Total	24	20	32	76

Conclusions and Interpretations

Our sample consisted of a mixed group of science and non science students on upper secondary level, who all had to study both science and non science subjects. Our data did not show a significant gender difference in motivation to learn science, neither for the whole sample nor for only the science students. This confirms on the one hand the results of Glynn and colleagues, who did not find such a difference either. But on the other hand it does not support their interpretation in terms of special learning conditions in non science classes, since our students learned under quite various conditions, depending of their favourite education subject. Our data suggest that a general gender difference in motivation to learn science does not exist.

There was however a highly significant positive correlation between the motivation to learn science and the braintype. This means that the more a student shows a systemizing cognitive style, the more he is motivated to learn science in general.

This is not really surprising, given the definition of a systemizing cognitive style as a drive and ability to analyze the rules underlying a system, in order to predict its behaviour. The remarkable aspect of this point however is that the braintype is a stable attribute of personality and that that there is a valid and reliable way to test it. Our data support our research hypothesis that the braintype is a more important factor for the motivation to learn science than gender.

To explain why this is so, it should be realized that braintype and gender are not independent. Our data show that on average female students are empathizers whilst male students are systemizers. This difference is highly significant and also supported by a large body of previous research which confirms this to be a general gender pattern.

We therefore propose to interpret the correlation between braintype and motivation towards learning science in terms of a first order relation, whilst the gender-motivation-difference (which can be seen in the data but is not significant) is only a second order phenomenon: Because men are generally more systemizing than women, and systemizers have a higher motivation to learn science than empathizers, male students tend to have a higher SMQ than female students.

If it is true that the braintype is a first order aspect influencing the SMQ, this could also explain our findings, that our science students on average are significantly higher systemisers than our non science students, and that the braintype, but not sex, predicts if these students join a science class or a non science class. These findings are consistent with previous research results, which showed that physical science degree students scored significantly lower on the EQ and significantly higher on the SQ and suggested, that the academic subject one ends up studying may be better predicted by one's cognitive style than by one's sex.

Since to join a science resp. a non science class obviously indicates a more resp. less positive overall attitude to learning science, then our results also suggest that the empirical evidence of male students having consistently more positive attitudes to school science than female students (as quoted in the introduction of this article) also is a second order phenomenon, and that the "basic player" in the background again is the braintype. This would mean that it is not the male students that have a more positive attitude towards science in general but the systemisers – and because boys generally are more systemizing than girls, their attitude is more positive towards science, especially towards physics, the school subject that probably most favours a systemizing approach of nature.

Of course our data cannot verify this on face value intriguing hypothesis. More research should be done to cast more light on that issue. Because attitudes – being a very heterogeneous body of different aspects – are more difficult to quantify than motivation, it could be appropriate to use a mixed approach to this research, for example by measuring the braintype by means of standardized questionnaires and exploring attitudes by a qualitative methodology.

We see a possible flaw in our data as to the high average of our students' SMQ. We explain this unexpected result by taking into consideration that these students have been tested in a science learning centre, where they had been attending a specially designed module on medical genetics. These classes could therefore have felt inspired by the course. It could also be that these classes hat consented to visit the science centre because of their intrinsic high motivation towards science, or their science teachers could be highly engaged in his/her teaching. This context could have biased our results and should be avoided in future research on this topic.

As we remarked at the beginning of this article, our research is meant as piece to the puzzle. More research must be done to be able to reliably link these findings to the situation in the science class room. Nevertheless, we would like to end this article – with due precaution – by outlining some thoughts emerging as possible implication for school science. Our results seem to point out that good systemizers have a high motivation towards science. Going back to the definition of Glynn and colleagues, this is "the internal state that arouses, directs, and sustains students' behavior toward achieving certain goals." Good systemizers are therefore not *per definition* good in (school) science, but they strive for it, which makes a good start to become a successful science student.

The challenge for school science seems to be – at least from the point of view of our results – the students with an empathizing braintype. It could be an interesting research question, how they should be approached to improve the systemizing dimension of their cognitive style, i.e. their " drive and ability to analyze the rules underlying a system, in order to predict its behaviour." Our findings suggest that a success in improving the systemizing dimension of their cognitive style could spontaneously entail an improvement of their motivation towards science. More research must show, if and to what degree the biological amount of systemizing can be improved and how this could be done.

References

Bandura, A. (2001). Social cognitive theory: An agentive perspective. *Annual Review of Psychology, 52*, 1-26.

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences, 6*, 248-254.

Baron-Cohen, S. (2004). *Vom ersten Tag an anders. Das weibliche und das männliche Gehirn.* Düsseldorf und Zürich: Walter Verlag.

Baron-Cohen, S., Knickmeyer, R., & Belmonte, M. K. (2005). Sex Differences in the Brain: Implications for Explaining Autism. *Science 310*(819-823).

Baron-Cohen, S., Richler, J., Bisarya, D., Gurunathan, N., & Wheelwright, S. (2003). The systemizing quotient: an investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences *Philosophical Transactions: Biological Sciences, 358*(1430), 361-374.

Baron-Cohen, S., & Wheelwright, S. (2004). The empathy quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders, 34*, 163-175.

Billington, J., Baron-Cohen, S., & Wheelwright, S. (2007). Cognitive style predicts entry into physical sciences and humanities: Questionnaire and performance tests of empathy and systemizing. *Learning and Individual Differences, 17*, 260-268.

Davis, M. H. (1980). A multidimensional approach to individual diferences in empathy. *JSAS Catalogue of Selected Documents in Psychology, 10*, 85.

Eisenberg, N. (2002). Empathy-related emotional responses, altruism, and their socialization. In R. J. Davidson & A. Harrington (Eds.), *Visions of Compassion: Western scientists and Tibetan Buddhists examine human nature.* London: Oxford University Press.

Glynn, S. M., & Koballa, T. R. (2006). Motivation to learn in college science. In M. J. & L. W. H. (Eds.), Handbook of college science teaching. (pp. 25-32). Arlington, VA: National Science Teachers Association Press.

Glynn, S. M., Taasoobshirazi, G., & Brickman, P. (2007). Nonscience Majors Learning Science: A Theoretical Model of Motivation. Journal of Research in Science Teaching, 44(8), 1088-1107.

Green, B. A., & DeBacker, T. K. (2004). gender and orientations toward the future: Links to motivation. Educational Psychology Review, 16, 91-120.

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: a review of the literature and its implications. international Jounal of Science Education, 25(9), 1049-1079.

Wellman, H. (1990). The Child's Theory of Mind. Cambridge, MA: MIT Press.

Wheelwright, S., Baron-Cohen, S., Goldenfeld, N., Delaney, J., Fine, D., Smith, R., et al. (2006a). Predicting Autism Spectrum Quotient (AQ) from the Systemizing Quotient-Revised (SQ-R) and Empathy Quotient (EQ). Brain Research, 1079, 47-46.

Wheelwright, S., Baron-Cohen, S., Goldenfeld, N., Delaney, J., Fine, D., Smith, R., et al. (2006b). Predicting Autism Spectrum Quotient (AQ) from the Systemizing Quotient-Revised (SQ-R) and Quotient (EQ). Brain Research, 1079, 47-56.
Witkin, H. A., Dyk, R., Faterson, H. F., Goodenough, D. G., & Karp, S. A. (1962). Personality through perception. New York: Harper & Row.

Appendix

Science Motivation Questionnaire (SMQ) (Glynn et al., 2007)
In order to better understand what you think and feel about your college science courses, please respond to each of the following statements from the perspective of: "When I am in a college science course . . ."
01. I enjoy learning the science. * Never * Rarely * Sometimes * Usually * Always
02. The science I learn relates to my personal goals. * Never * Rarely * Sometimes * Usually * Always
03. I like to do better than the other students on the science tests. * Never * Rarely * Sometimes * Usually * Always
04. I am nervous about how I will do on the science tests. * Never * Rarely * Sometimes * Usually * Always
05. If I am having trouble learning the science, I try to figure out why. * Never * Rarely * Sometimes * Usually * Always
06. I become anxious when it is time to take a science test. * Never * Rarely * Sometimes * Usually * Always
07. Earning a good science grade is important to me. * Never * Rarely * Sometimes * Usually * Always
08. I put enough effort into learning the science. * Never * Rarely * Sometimes * Usually * Always
09. I use strategies that ensure I learn the science well. * Never * Rarely * Sometimes * Usually * Always
10. I think about how learning the science can help me get a good job. * Never * Rarely * Sometimes * Usually * Always
11. I think about how the science I learn will be helpful to me. * Never * Rarely * Sometimes * Usually * Always
12. I expect to do as well as or better than other students in the science course. * Never * Rarely * Sometimes * Usually * Always
13. I worry about failing the science tests. * Never * Rarely * Sometimes * Usually * Always
14. I am concerned that the other students are better in science. * Never * Rarely * Sometimes * Usually * Always
15. I think about how my science grade will affect my overall grade point average. * Never * Rarely * Sometimes * Usually * Always
16. The science I learn is more important to me than the grade I receive. * Never * Rarely * Sometimes * Usually * Always
17. I think about how learning the science can help my career. * Never * Rarely * Sometimes * Usually * Always
18. I hate taking the science tests. * Never * Rarely * Sometimes * Usually * Always
19. I think about how I will use the science I learn. * Never * Rarely * Sometimes * Usually * Always
20. It is my fault, if I do not understand the science. * Never * Rarely * Sometimes * Usually * Always
21. I am confident I will do well on the science labs and projects. * Never * Rarely * Sometimes * Usually * Always
22. I find learning the science interesting. * Never * Rarely * Sometimes * Usually * Always
23. The science I learn is relevant to my life. * Never * Rarely * Sometimes * Usually * Always
24. I believe I can master the knowledge and skills in the science course. * Never * Rarely * Sometimes * Usually * Always
25. The science I learn has practical value for me. * Never * Rarely * Sometimes * Usually * Always
26. I prepare well for the science tests and labs. * Never * Rarely * Sometimes * Usually * Always
27. I like science that challenges me. * Never * Rarely * Sometimes * Usually * Always
28. I am confident I will do well on the science tests. * Never * Rarely * Sometimes * Usually * Always
29. I believe I can earn a grade of "A" in the science course. * Never * Rarely * Sometimes * Usually * Always
30. Understanding the science gives me a sense of accomplishment. * Never * Rarely * Sometimes * Usually * Always

CONCEPTIONS OF FINNISH 7-8 YEARS OLD PUPILS ON HUMAN ANATOMY AND PHYSIOLOGY – A PHENOMENOGRAPHIC CASE STUDY

Eila Jeronen
University of Oulu

Heli Niemitalo
University of Oulu

Juha Jeronen
University of Jyväskylä

Riitta-Liisa Korkeamäki
University of Oulu

Abstract

What do young pupils know of what is inside them and how their organs function? The topic is important concerning biology education because conceptions constructed during the primary level have effects on scientific thinking and its development later on. The topic has been studied broadly in other countries, however its investigation has not been popular in Finland. The study is a qualitative case study. The data was gathered in spring 2008. It includes drawings and interviews of 20 pupils from the 1st form at primary level. Its methodological basis is phenomenographic. The resultant data were analyzed qualitatively and quantitatively. Analysis shows that some of the pupils begun to form scientific conceptions of the structures and the functions of the organs in the human body. Some of them had little appreciation, including only every-day concepts, and some had no conceptions at all. Misconceptions were common.

Introduction

In the study, we describe young Finnish pupils' understandings of their own internal structures and their functions. This is a part of an on-going study in the Nordic countries. The Finnish data analyzed here was chosen partly because it was the first one we got and partly because it was identified by us as a good, representative example. 7–8 years-old primary level pupils (further PLP) were asked to draw what they thought was inside themselves. Thereafter they were interviewed. We chose PLP on the grounds that in just about every Nordic country the great majority of seven-eight year-olds have begun school.

Already pre-school children know that the brain is for thinking and that it is located in the head. Children of 10 years-old and above understand that the brain regulates physiological as well as emotional reactions (Johnson & Wellman, 1982, 223–229). The organs best described by children are bones and the digestive and respiratory systems, and the worst ones are muscles, the endocrine system and blood circulation (Reiss & Tunnicliffe, 2002). Small children know the location of the heart, its structure and function but the location of the bones, the lungs and the stomach are not known although children know their meaning. They are also not familiar with blood circulation and blood. The location of the brain is well drawn (Óskarsdottír, 2006).

Learning processes and research results depend not only on guidelines and questions given in the investigation situation but also on a cultural basis. Reiss and Tunnicliffe (2002, 58–61) have noticed that children from Brazil, Northern Ireland, and Taiwan know more of the structure and function of the human body than the children from other countries. E.g. in Taiwan, seven years-old boys made better drawings than same-aged girls. They could not notice these kinds of differences between boys and girls in other countries.

This study is important for the development of Finnish biology education in primary schools. We are unaware of any work from Finland that systematically, simultaneously qualitatively and quantitatively examines the following issues, which are focused on in this study:

1) What kinds of conceptions, as revealed by drawings and interviews, do Finnish PLP have of their body before teaching about the human body begins?

2) How does such knowledge of boys (further B) and girls (further G) differ between the various human organs and organ systems?

Material and Methods

This qualitative case study involved a total of 20 Finnish PLP (12 B; 8 G) from a school in Northern Finland. The school was familiar to an author from her earlier contacts and the teachers and pupils were willing to participate in the research voluntarily. The parents of the pupils had given their permission for the participation of their children. The data was gathered in spring 2008. At the beginning, the pupils were asked to draw what they thought was inside them. They were given 15 minutes to complete their drawings. Drawings were used because they are particularly suitable for international studies. Children also like to draw and they try to do as exact picture as possible based on their knowledge.

When the drawings were ready, the pupils were interviewed based on their drawings. If drawings are used alone, there is a risk that children copy pictures from their textbooks and drawings do not present their own conceptions (Óskarsdottír, 2006, 136). When using interviews beside drawings, the method offers an opportunity to avoid interpretation mistakes during the analysis phase (Osborne et al., 1992, 29).

The methodological basis of the study is phenomenographic (Niikko, 2003). The resultant data were analyzed qualitatively and quantitatively (Greene et al., 1989). The drawings were sorted out, attempting to arrange them in a ranked order which the authors felt reflected different levels of biological understanding (Óskarsdottír, 2006) (Table 1). Statistical tests were not made because of the small number of participants.

Table 1. Knowledge scales (Óskarsdottír, 2006).

Bones – skeleton	
Level 1	No bones.
Level 2	Bones indicated by simple lines or circles.
Level 3	Bones indicated by 'dog bone shape' and at random or throughout body.
Level 4	One type of bone in its appropriate position.
Level 5	At least two types of bone (e.g. backbone and ribs) indicated in their appropriate position.
Level 6	Definite vertebrate skeletal organization shown (i.e. backbone, skull and limbs and/or ribs).
Level 7	Comprehensive skeleton (i.e. connections between backbone, skull, ribs and limbs).

Organs
Level 1 No representation of internal structure.
Level 2 One internal organ (e.g. brain or heart) placed at random.
Level 3 One internal organ (e.g. brain or heart) in appropriate position.
Level 4 Two internal organs (e.g. brain, heart or stomach) placed at random.
Level 5 Two internal organs (e.g. brain, heart or stomach) in appropriate positions but no extensive relationships indicated between them.
Level 6 More than two internal organs in appropriate position but no extensive relationships indicated between them.
Level 7 More than two internal organs in appropriate position and one organ system indicated (e.g. gut connecting head to anus or connections between heart and blood vessels).
Level 8 Two or more major organ systems indicated out of digestive, circulatory, gaseous exchange and nervous systems.

Results and discussion

The drawing analysis suggests that the boys had better knowledge level of the bones than the girls (B/G: levels 1–3: 33/63%; level 4: 17/37; levels 5–6: 50/0 %). For some examples, see Figure 1. Both the boys and the girls knew hand and foot bones the best (B/G: 67/25 %) and ribs (16/0 %), backbone (8/0 %), and toe bones (0/13 %) the worst. The skull was mentioned by five boys and a girl. However its drawing was problematic nearly for all of them. Finger bones were drawn by two boys and a girl. Connections between backbone, skull, ribs, and limbs were unknown. The autonomous descriptions of bones included expressions such as “Bones are hard and white.” and “The bones support and move the body.” A pupil stated that the body shape is based on the bones. About half of the pupils described muscles saying that they are in their hands and feet. They also knew that human beings need muscles for moving and exerting force. The result supports the findings of Osborne et al. (1992, 35).

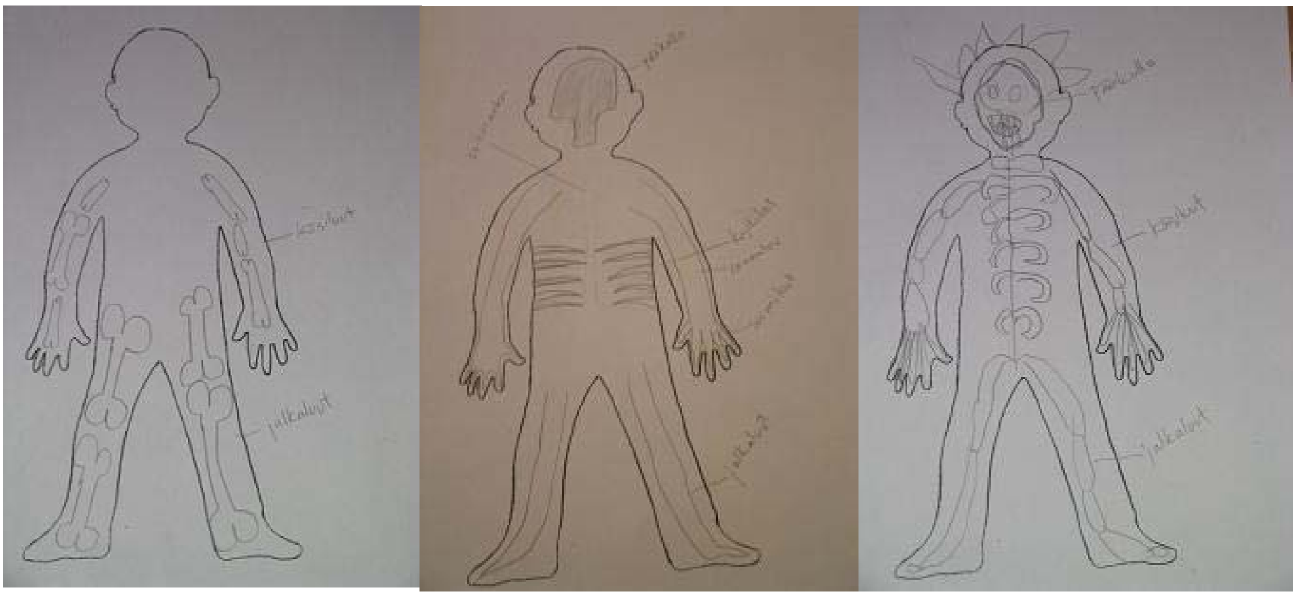


Figure 1. Some examples of bone drawings. Left, Level 3. Middle, Level 5. Right, Level 6.

The boys knew the human internal structure better than the girls (B/G: levels 1–3: 25/37%; level 4: 0/13; levels 5–6: 75/50 %). See Figure 2 for some examples. Both groups knew very well the location of the heart (10 B/7 G). Over half of them mentioned that “The heart bounces and keeps a human being alive.” Two pupils knew that the heart pumps blood into the vessels. Three pupils stated that the heart pumps oxygen. A girl said that the heart makes blood. Other pupils did not know the function of the heart at all. One pupil did not know anything of the heart. The red blood was drawn by some pupils all over the body and some of them knew that it circulates in vessels. One pupil drew both situations. The result resembles the findings of Óskarsdottir (2006).

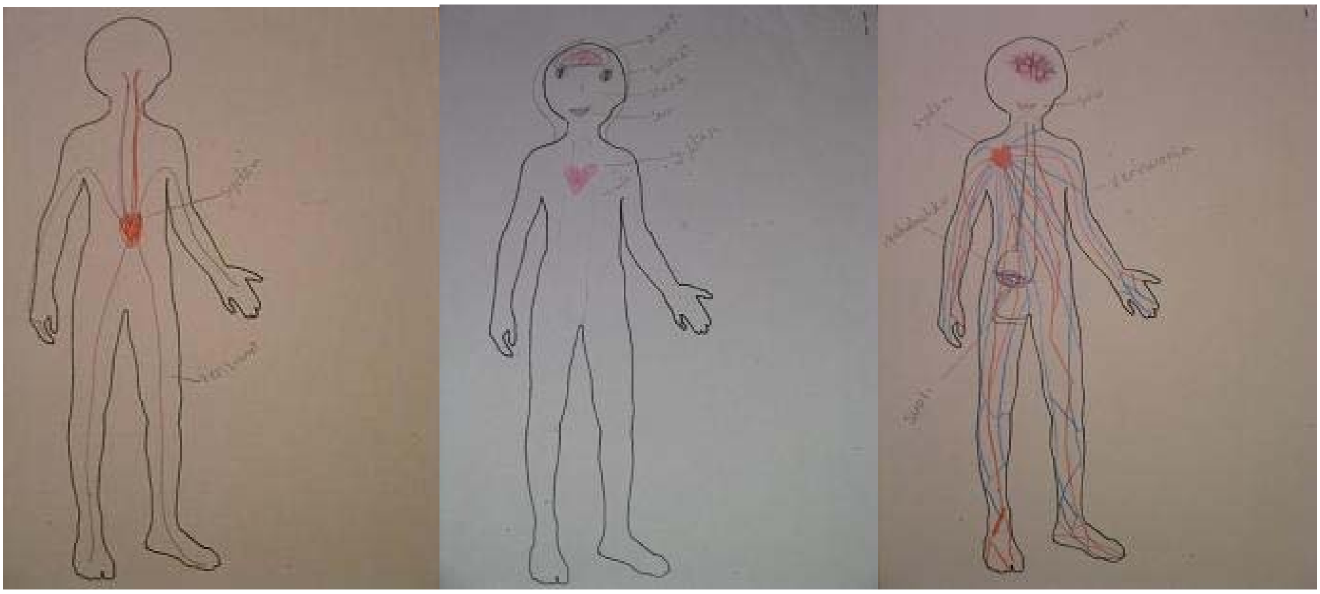


Figure 2. Some examples of internal organs drawings. Left, Level 2. Middle, Level 5. Right, Level 8.

Ten boys knew the brain very well and also three girls mentioned them. The result supports the findings of Óskarsdottir (2006). Most of the pupils told that human beings need the brain for remembering, knowing, thinking, and understanding. Two boys thought that the brain orders a person to do something. Emotions were not connected with the brain functions (cf. Johnson & Wellman, 1982, 223). Lungs were known by a boy and two girls. However, they mixed hearth and lung functions with each other. The result supports the findings of Osborne et al. (1992, 36–42) but not the findings of Óskarsdottír (2006).

Most of the pupils were able to tell about digestion although they could not draw the digestive system (cf. Osborne et al., 1992, 36–40). Six pupils knew that food goes from the mouth to the stomach. Four of them were able to connect the mouth, the stomach and the intestine with each other and two told about feces. Although, some of the pupils knew well the names and functions of the digestive system, there were also many misconceptions. The names of the digestive system parts were mixed or they were not known at all. The kidneys and the liver were mentioned only by a boy. The reproduction system was not described at all.

The reliability of interpretation of the results was increased by interviewing all respondents. The researcher was familiar to the pupils beforehand, creating a friendly atmosphere for interaction. The drawings supported the interviews in categorization of the pupils' understanding, making the categorization more reliable. This categorization of the data was done by two independent researchers. The results of the study were compared to existing results in the literature. However, it should be noted that this was a case study, and the results cannot be generalized due to the small number of respondents (n=20).

The fact that primary level pupils who had not yet had teaching of human body had some understanding concerning it reflects the public awareness and usage of terms related to the internal organization of the body in every-day life. Some of the pupils had also started to construct scientific conceptions of the human body. It will be interesting to see whether studies in other Northern countries will produce similar or different results. Learning about human beings begins at home and reflect the public understanding in this area. However, teachers have a crucial role when developing pupils' biological conceptions.

References

Greene, J.C., Caracelli, V.J. & Graham, W.F. (1989). "Toward a conceptual framework for mixed-method evaluation design." Educational Evaluation and Policy Analysis, 11(3), 255–74.

Johnson, C.N. & Wellman, H.M. (1982). Children's developing conceptions of the mind and brain. Child Development, 53(1), 222–234.

Niikko, A. (2003). Fenomenografia kasvatustieteellisessä tutkimuksessa. Joensuun yliopisto. Kasvatustieteiden tiedekunnan tutkimuksia, 85.

Osborne, J., Wadsworth, P. & Black, P. (1992). Processes of Life: Primary space project research report. Liverpool: Liverpool University Press.

Óskarsdottír, G. (2006). Children's ideas about the human body – How childern's drawings can be used as part of assessment. In P. Nielsen, (Ed.) Innovation og aflaering. 2. bog. Den 9. nordiske laereruddannelseskongres. Färöerna. 130–137.

Reiss, M. & Tunnicliffe, S.D. (2002). An international study of young peoples' drawings of what is inside themselves. Journal of Biological Education, 36(2), 58–64.

STUDENTS' UNDERSTANDING OF THE NATURAL WORLD: HOW DO SIXTH GRADE STUDENTS PERCEIVE THE FLOWER?

Şirin İlkörücü Göçmençelebi
Uludag University

Abstract

The aim of this study is to investigate students' perception of the concept of a flower using their drawings. Secondary school students were firstly asked to draw a flower and then to label the component parts. All drawings were collected, evaluated and grouped based on the labeling_and the actual drawing. This study classified two main groups of drawings: the concept of the parts of a flower (comprehensive, partial, incorrect) and the drawing of the structure of a flower (visual-iconic and redundant). In this investigation it was found that only 39.9% of the students drew and correctly identified the structure of the flower. In addition, it was found that 51.0% of students were not able to draw the flower correctly. This study reveals how students expressed their mental models of flowers through their drawings

Introduction

In science education, learning concepts correctly is important to prevent students from misinterpreting natural phenomena. Teachers need effective ways to understand students' understanding of science concepts correctly. There are many ways of gathering information (interview, diagnostic test etc.) about students' understanding of scientific perception (Jewell, 2002; Lin 2004; Hellden 2004). However, instructors can gather large amounts of data on the mental models that students have about scientific concepts using simple drawings and thus improve the teaching and learning process (Hoese & Casem 2007; Köse 2008). As Tunnicliffe and Moussouri (2003) stated, the great majority of the methods that have mostly been used by science educators rely on students being able to talk about or write about science. Among these, the drawings of students are a non-verbal way of making and communicating meaning and can represent children's thoughts.

Tunnicliffe and Reiss (2000) indicated that school children aged 5-14 years use predominantly anatomical reasons (leaf, flower/fruit, stem, form of growth, dimension etc.) when identifying plants. Gattt et al. (2007) stated that young children regard a plant as being something that is small and green with a stalk and leaves. Many studies indicate that in general students have some misconceptions about flowers or flowering plants (Jewell 2002; Hellden, 2004; Lin 2004; Gatt et all. 2007). This study was to investigate the students' perception of the concept of a flower using their drawings. Students were firstly asked to draw a flower and then to label the component parts.

Rationale

The flowering plant unit is regularly taught in Grade 6 in Turkish schools. The place of flowering plants in the Turkish National Curriculum is to help students understand the natural world which can be seen easily in everyday life. Although the concept of a flower has been seen as a simple part of flowering plants, students have some misconceptions. The flower relates to the other major life cycle processes of flowering plants such as fertilization of plants which is important for fruit and helps to protect and disperse the seeds. Students' everyday experiences tend to dominate scientific knowledge that is taught during lessons. Students' drawings of flowers with which they are familiar, enable educators to make interpretations into clear categories. Using a single category to evaluate students' drawings will help educators to understand students' mental perceptions, in particular, their misconceptions of the natural world related to their learning at school.

Methods

This research was carried out in three state schools in Bursa, Turkey. 143 pupils from Grade 6 class (11 and 12 years old) participated in this study. The sample included 65 boys and 78 girls. The research data was collected over 2 years from 48 students in the first year and 95 students in the second year. In all cases the data was collected from the students within 2 weeks of the information having been taught in a lesson.

Data analysis: Drawings collected from the students were classified into two main groups; the concepts of the parts of a flower and the drawing of the structure of a flower. The evaluation of the labeling was made according to the parts which had been taught.

Examining *the concept* of part of a flower, three categories were defined as based on previous studies (Hoese & Casem, 2007; Köse, 2008) as given below;

a) *Comprehensive labelling:* All the parts of the flower labelled correctly.

b) *Partial labelling:* Some parts of the flower labelled incorrectly.

c) *Incorrect labelling:* All parts of flower labelled incorrectly or not labelled.

Examining *the drawings* of the structure of the flower, two different categories were defined as based on previous studies (Hoese & Casem, 2007; Köse, 2008) as given below;

V*isual-iconic drawing:* Structure of the flower was visually drawn correctly as a reproduction organ of flowering plant. Internal structure of flower not considered.

Redundant drawing: In addition to the structure of the flower, flowering plants were visually drawn with extra parts.

One of the aims of this research was to investigate students' perception of a flower as distinct from a flowering plant. Therefore, if the student drew only the flower even if there were missing parts of the actual flower, it was categorized visual/iconic. If the student drew a flower with extra parts (leaves, roots etc.) it was categorized as redundant drawing.

Results

The results are shown in Tables 1 and 2. 39.9% of the students drew and correctly identified the component parts of the flower. Of these 57 students, 19.3% drew extra parts. 32.2% of the total number of students made a correct drawing with incorrect labeling. Of these 46 students, 52.2% drew extra parts or had parts missing. 28% of the total number of students made incorrect drawing and of those, 95% drew extra parts or had parts missing. 51% of the students were not able to draw the flower correctly (Table 2). 60.2 % were not able to label the component parts (Table 1).

Table 1. Percentages of concept categories by school

Schools	comprehensive labelling	partial labelling	incorrect labelling
School A (valid *n* = 48)	16.1% (*n* =23)	11.2% (*n* =16)	6.3 %(*n* = 9)
School B (valid *n* = 42)	11.2% (*n* =16)	11.2% (*n* =16)	7.0 % (*n* = 10)
School C (valid *n* = 53)	12.6% (*n* =18)	9.8% (*n* =14)	14.7 % (*n* = 21)
Total	39.9% (*n* =57)	32.2 % (*n* = 46)	28.0% (*n* =40)

Table 2 Percentages of related drawing groups according to concept categories

	visual-iconic drawing	redundant drawing
comprehensive labelling (valid *n* = 57)	80.7% (*n* =46)	19.3% (*n* =11)
partial labelling (valid *n* = 46)	47.8% (*n* =22)	52.2% (*n* =24)
incorrect labelling (valid *n* = 40)	5.0% (*n* =2)	95.0% (*n* =38)
Total	49.0% (*n* =70)	51.0 % (*n* = 73)

Conclusions and Implications

Students expressed their mental models of flowers through their drawing. Students have difficulty in perceiving the flower as a reproduction organ of a flowering plant. The results of this study indicate that approximately 50% of students perceive the flower as a flowering plant. This misconception most probably arises because of the concept of flowers in daily life. Focusing on a particular set of structures will encourage the students to have a more comparative approach to flowering plant observation.

It seems that students who correctly comprehend the concept, also perceive the flower concept correctly (80.7%) as a part of a flowering plant in their drawings. Thus showing that there may be a relationship between conceptual perception and iconic-visual perception. However, we need to understand whether the students have fully understood or only memorized. For this reason, drawing activities should be used in conjunction with interviews to explore students' ideas about their daily world and academic concepts.

References

Gatt, S., Tunnicliffe S. D., Borg, K. & Lautier, K. (2007). Young Maltese Children's Ideas About Plants. *Journal of Biological Education*, 41(3): 117-121.

Hellden, G. F. (2004). A Study of Recurring Core Developmental Features in Student's Conceptions of Some Key Ecological Process. *Canadian Journal of Science, Mathematics & Technology,* 4(1): 59-70.

Hoese, W. J.& Casem, M. L. (2007). Drawing out misconceptions: Assessing Student Mental Models in Biology. Retrived January 4, 2009 from www.bioliteracy.net/Readings/papersSubmittedPDF/Hoese%20and%20Casem.pdf

Köse, S. (2008). Diagnosing Student Misconceptions: Using Drawings as a Research Method. *World Applied Science Journa*l, 3(2): 283-293.

Lın S.W. (2004). Development and Application of a Two Tier Diagnostic Test for High School Students' Understanding of Flowering Plant Growth and Development. *International Journal of Science and Mathematics Education*, 2:175-199.

Jewell N. (2002) Examining Children's Models of Seed. *Journal of Biological Education*, 36(3): 116-122.

Tunnicliffe, S. D.& Reiss M. J. (2000). Building a Model of the Environment: How do Children see plants? *Journal of Biological Education* 34(4):172-177.

Tunnicliffe S. D. & Moussouri, T. (2003). Methods for Assessing Out of School Science Learning Experiences. Paper presented at the Annual Meeting ESERA. http://www1.phys.uu.nl/esera2003/programme/authors.htm

EFFECT OF USING NEWSPAPER CLIPPINGS IN SCIENCE AND TECHNOLOGY COURSE ON STUDENTS' ATTITUDES TOWARDS SCIENCE

Esma Buluş Kırıkkaya
Kocaeli Üniversitesi

Şebnem İşeri
Kocaeli Üniversitesi

Abstract

Informal education sources are tools which simplify connection with daily life. Informal atmosphere that formed by using informal education tools like newspaper, internet and television in classroom will provide students to develop positive attitude. Through these tools newspaper is a material that is economic, easy to find and also can be used in class easily. Making use of newspaper activities especially in science courses will affect students' attitudes towards science positively. To this end, 4th grade students were included to the study from similar schools that selected through primary schools in Kocaeli. Study lasted along "Introduction to Matter" unit. Suitable acquisitions in unit were selected and 10 activities were arranged. Activities were provided to last right behind to the related topics by performing minimum one activity per week. For determining variation in students' attitudes a pre-test and after performing all activities a post-test were applied to experiment and control groups. When experiment and control groups' post test mean scores checked a significance difference has been found in favor of experiment group.
Keywords: *Newspaper activities, attitude towards science*

Introduction

Media tools are pretty affective in adult education. Especially media tools like television, internet and newspaper are important informal education environments. However, these tools have contents which appeals to people of all ages and they are powerful enough to provide intended changes in peoples behaviors. Science information, which has been got from informal sources, is important because it is related to up to date information and daily life (Halkia& Mantzouridis, 2005).

Above all remembered experiences about science are generally gained in informal environments (NSF, 1998). In a part of his study, Elliott (2006) gave a list of biotechnical words to student-teachers and wanted them to sign the words that they heard before and to explain what they understood from these words. According to student-teachers' responses, it was assumed that their pre-knowledge about the topic would derive from exposure to media by researcher. If experiences gained in informal environments will transfer to class; education can be more permanent and effective. Transferring publications like newspaper and technologies like internet and television to classroom provide students to gain informal experiences in a formal environment. Although more schools have technologies like television and internet, even in these schools it is impossible to using them in all classes. Also newspapers provide students to make observations, ask questions, investigate the information and discuss their views (Deveci, 2005). So it is thought that using a material like newspaper in science courses which is cheap and easy to find would be more favorable.

Using newspaper in course will simplify making connections between science and daily life. Jarman and Mc Clune (2002) investigated how teachers were using newspapers in science course. According to this study's results most of the teachers are using newspapers for presentation and as a source. And the using reasons are generally connecting to daily life and improving interest to science. Ivan and Shibley (2003) found in their study that discussing daily discoveries about science which mentioned in newspaper attract students' attention and scientific articles simplifies to understand nature of science.

Using this material in class will effect students' achievements positively. NIE (Newspaper in Education) is an extensive program in abroad that newspapers and schools included for years. With this program it is intended to using newspaper in courses and results of applications were analyzed. In a NIE report prepared by Sullivan (2002) it is mentioned that using newspapers in courses effects student achievement about 10%.

The reasons of teachers doesn't use newspaper in classroom are they didn't receive training and they couldn't find any source which explains how to use it. Vockell and Cusick (1995) determined that some of the teachers who worked in non-NIE schools hesitated to use newspaper. Also most of the teachers put forwards reasons like newspapers didn't suitable to the curriculum; they didn't receive training and lack of time. On the other side the reasons of using newspaper were shown as giving students the opportunity of practicing the abilities that they gained in class, motivating them, containing more actual information than course books and providing information about different topics. Also Kachan, Guilbert and Bisanz (2005) found that while most of the biology teachers use newspaper frequently, only half of science teachers use it. In this study most of the teachers explained that they used newspaper for interrelating to STSE (Science, Technology, Society and Environment), for supporting curriculum and as discussing material.

Rationale

As a common media tool newspaper is also an effective informal education material. With its ability of providing connections to daily life, newspaper will be an important tool for Science and Technology education. The newspaper activities that students do in Science and Technology course will increase attendance of students to the course by forming an informal atmosphere in classroom and will help to change their attitude towards this course positively. In abroad newspaper used especially in study of languages, also in other courses it usage is being seen and there are studies that can guide teachers. The directive of TTKB (Board of Education) which was the numbered 3 decree that decided on 2008 about "Utilizing newspaper clippings in education process of primary education programs" is an important step on this subject. But it makes difficult to use this material that there isn't a guide for teachers which explains them how to use newspaper and contains activity samples they can use in course. Preparing a guide like this will increase the usage rate of newspapers in Science and Technology courses. In accordance to this, in this study it is intended to determine how the activities prepared with newspapers is effective in Science and Technology course. To this end it was investigated that the effect of using newspaper activities in Science and technology course on students' attitude towards science, the effect of gender on attitudes and students thoughts about prepared activities.

Problem Statements and Sub problems:

1. What is the effect of newspaper activities that used in Science and Technology course on students' attitudes towards science?
    1.1. What is the effect of using newspaper activities in Science and Technology courses on students' attitudes towards science according to gender?
2. What are the students' thoughts about prepared newspaper activities?

Methods

This study is semi-experimental and pre test-post test control group design was used. Along the study the newspaper activities performed with experiment group and with control group traditional methods and activities used. Units in the 4th grade's curriculum investigated and the unit which has the most appropriate time and most appropriate acquisitions was selected. According to these; second unit: 'Introduction to Matter' was included to study and activities lasted along this unit for 9 weeks.

Study Universe and Sample

Study universe consists of primary schools in Kocaeli and the sample randomly selected from similar primary schools. A science attitude pretest applied to 4th grade students and similar classes were selected according to students' attitude level. An experiment and a control group from two schools, totally 108 students from 4 classes included to study.

Activities

Selected unit contains 46 acquisitions and all of the acquisitions were investigated whether they were appropriate for newspaper activities. While preparing newspaper activities, researchers benefit from 2007 NIE Week Teachers Guide, 2008 NIE Week Teachers Guide and Sanderson's (2007) book about using newspapers in classroom. For the activities, 22 acquisitions were selected and 10 activities were prepared. Teachers' thoughts about relevance of activities and acquisitions were taken before the application. Also some experts investigated the activities for validity and reliability analysis. Every activity contains directives to teachers and student activity sheet. In the cause of class teacher executed the activities, information about application was given to teacher before every activity. Researchers observed the lessons which activities done.

Figure 1. Students are looking for the matters.

Activity 1: Matter Scavenger Hunt

The aim of this activity is; providing students to describe matters with the specialties that can be realized by five sense organs. Whole of the newspaper was used for this activity and students were wanted to find out matters that belong to given adjectives like transparent, opaque, soft and hard.

Activity 2: Find the Correct One

The aim of this activity is; providing students to use 'matter', 'object', 'material' and 'tool' conceptions correctly in a sentence. Whole of the newspaper was used in this activity and students were wanted to find five sentences which are includes 'matter', 'object', 'material' and 'tool' conceptions. After they found the sentences, it was asked that if these conceptions were used correctly or not.

Figure 2. Students are searching the sentences.

Figure 3. Students are reading the selected article.

Activity 3: Which Matter it was Made with?

The aim of this activity is; providing students to relate specialties of the matter and its usage in daily life. News articles were used in this activity. An article was selected about a material and its specialties about it before the activity. The article was read loudly by teacher. Students were wanted to read article and guess which matter the mentioned material could be made with.

Activity 4: Horrifying Smell

Figure 4. Students are discussing about the diffusion of gases.

The aim of this study is; providing students to comprehend that 'gases can diffuse in the atmosphere which it is in'. News articles were used in this activity. Students were wanted to read the article which was about sensed smell of natural gas in a big field. After they read it, it was asked that what the reason of sensing the smell of gas in a big field was.

Activity 5: What is the Physical Condition of the Matter?

Figure 5. Students are marking the matters.

The aim of this activity is; providing students to classifying the matter according to its condition. Whole of the newspaper was used for this activity. Students were wanted to find a symbol for solid, liquid and gas conditions of the matter. They found out matters from pictures and articles, then they marked matters in the newspaper with their symbols. Students were asked why they choose the symbols that they used and compared the markings with their classmates.

Activity 6: Find the Matter

Figure 6. Students are writing their answers.

The aim of this activity is; providing students to comprehend that which scales the matters could be measured with. Whole of the newspaper used in this activity. Students were wanted to find related matters to given categories like 'the matters which measure with balance scale'.

Figure 7. Students are reading the article.

Activity 7: Natural Resources are Exhausting

The aim of this activity is, providing students to explain why the natural resouces must be used carefully and the importance of informing people about this topic. News were used for this activity. A news was given to students about natural resources and students were wanted to answer questions according to news after they read it.

Activity 8: Fill in the Blanks

Figure 8. Students are comparing their answers.

The aim of this activity is; providing students to comprehend the transform between the conditions of matter. Whole of the newspaper was used in this activity. Before the activity some sentences were selected from the newspaper by teacher. Teacher read the sentences loudly, but without the words about the transform between conditions of matter. Students were wanted to find out the word which were not read and to write down the answer.

Figure 9. Students are preparing questions.

Activity 9: Let's Interview

The aim of this activity is; providing students to comprehend the seperation methods of mixtures. Students were wanted to prepare questions and to interview about making some of the foods which are prepared at home for winter. They found out the seperation method which was used for preparing the food. They were asked to write a news about the interview and the separation method.

Activity 10: Dear Teacher…

The aim of this study is, evaluating students and determining the difficulties they had during the unit. Students were wanted to write a letter to their teacher about the difficulties that they had along the unit. Then they exchanged the letters randomly and write answers to letters.

Figure 10. Students are writing letter.

Gathering and Analyzing Data

For determining the effect of newspaper activities on students' attitudes towards science, an attitude scale towards science which developed by Demirci was used. It consists of 32 items and it is a five point likert type scale. This scale's reliability was found .96 in Demirci's study (Cited in Özçelik, 2007). In this study reliability coefficient was found as .88.

Before the unit started for determining students initial attitude level a pre-test was applied. And for determining changing in students' attitudes a post-test was applied after unit finishes. Also student thoughts about activities were got. Hence, things that students done were reminded to them after all activities completed and they were wanted to write a paragraph about the activities they liked most and they liked last or they didn't like with the reasons. Analyses of the gained data were done with SPSS 11.5 packaged software.

Results

The findings about the students' attitudes and their thoughts about the activities were investigated at this part of the study.

In Table 1, experiment and control groups' pre and post test science attitude scores evaluated separately between groups. Experiment and control groups' pre-test scores compared and it was seen that there isn't a meaningful difference between the groups ($t=.928$, $p>.05$). And when results of the post-test of science attitude which applied after doing the newspaper activities compared it was seen that there is a meaningful difference in favor of experiment group ($t=2.027$, $p<.05$).

Table 1. Comparison of Experiment and Control Groups' Pre-test and Post-test Science Attitude Scores

	Group	N	Mean	Standard deviation	T	P
Pre-test	Experiment	51	4.24	.538	.928	.355
	Control	57	4.13	.649		
Post-test	Experiment	51	4.40	.549	2.027	.045
	Control	57	4.15	.735		

In Table 2, experiment group's pre-test and post-test science attitude scores and control group's pre-test and post test scores compared in groups. When experiment group's pre and post-test results investigated it is seen that a meaningful difference occurred in favor of post-test (t=2.104, p<.05). But a meaningful difference didn't occur in control group's science attitude level (t=.185, p>.05).

Table 2. Comparison of Experiment Group's Pre-Post test Science Attitude Scores and Control Group's Pre-Post test Science Attitude Scores

Group		N	Mean	Standard deviation	T	P
Experiment	Pre-test	51	4.24	.538	-2.104	.040
	Post-test	51	4.40	.549		
Control	Pre-test	57	4.13	.649	-.185	.854
	Post-test	57	4.15	.735		

In Table 3, it was investigated that if there was a difference between the attitude scores of experiment and control groups according to gender. According to this table, when variation of attitude investigated a significant difference was found in favour of girls in experiment group (t=2.590, p<.05) while there wasn't a difference according to gender in control group (t=1.544, p>.05). And in post-test results, there wasn't a significant difference according to gender for experiment group (t=1.493, p>.05), but it was seen that a significant difference occurred in control group (t=2.979, p<.05). While activities that were done with experiment group made up the difference between girls and boys science attitude, they increased the difference in favor of girls in control group.

Table 3. Variation of Science Attitude Occurs in Experiment and Control Groups According to Gender

Group		Gender	N	Mean	Standard deviation	T	P
Experiment	Pre-test	Girl	26	4.41	.502	2.509	.015
		Boy	25	4.05	.521		
	Post-test	Girl	26	4.51	.517	1.493	.142
		Boy	25	4.29	.568		
Control	Pre-test	Girl	31	4.25	.572	1.544	.128
		Boy	26	3.99	.715		
	Post-test	Girl	31	4.39	.564	2.979	.004
		Boy	26	3.85	.813		

Schoolgirls' and schoolboys' science attitude scores were investigated in Table 4. According to this table when the pre-test scores were compared it was seen that there wasn't a meaningful difference between mean score of girls in experiment group and in control group (t=1.135, p>.05), also there wasn't a meaningful difference between boys attitude scores (t=.381, p>05). When the post-test scores were investigated it was found that there wasn't any meaningful difference between schoolgirls' scores (t=.817, p>.05) but there was a meaningful difference in favor of schoolboys between boys in experiment and control groups (t=2.214, p<.05).

Table 4. Comparison of Schoolgirls' and Schoolboys' Pre-test-Post-test Science Attitude Scores

Gender			N	Mean	Standard deviation	T	P
Girl	Pre-test	Experiment	26	4.41	.502	1.135	.261
		Control	31	4.25	.572		
	Post-test	Experiment	26	4.51	.517	.817	.417
		Control	31	4.39	.564		
Boy	Pre-test	Experiment	25	4.05	.521	.381	.705
		Control	26	3.99	.715		
	Post-test	Experiment	25	4.29	.568	2.214	.032
		Control	26	3.85	.813		

While the mean score of schoolgirls in experiment was very similar to girls in control group, the mean scores of students in experiment and control groups were different. This situation could be arising because of schoolboys' mean scores made bigger difference than schoolgirls' scores.

Student thoughts about activities were generally positive. It was found that students liked all of the activities and the most important reason mentioned by students is finding activities entertaining. Students' thoughts about the activities they like are below with their own sentences:

"I liked 'Matter Scavenger Hunt', because it was easy."

"I liked 'Matter Scavenger Hunt', because it was funny."

"I liked 'Dear Teacher…' activity a lot, because it is enjoyable."

"I liked 'Dear Teacher'… activity a lot, because it is enjoyable and writing letter was good."

"I liked 'Horrifying Smell', because it was more beautiful than the others."

"I liked all of them, because they are very easy."

"I liked all of them, because they were most enjoyable activities in Science and Technology course."

"I liked all of them, because they are enjoyable and we used newspapers."

"I liked 'Dear Teacher…' activity, because I could tell my teacher (the difficulties that have about this topic)."

"I liked 'What is the Physical Condition of the Matter?' activity, because I enjoy finding from newspaper."

Very few of the students mentioned they didn't like some of the activities. Because they had difficulties with the activities or they didn't enjoy them. The reasons of disliking are given below with their own sentences:

"I didn't like 'Natural Resources are Exhausting', because I didn't enjoy it."

"When I couldn't do the activity papers that my teacher gave, I didn't like those activities."

"I didn't like 'Horrifying Smell', because I couldn't understand it."

"I didn't like 'Find the Correct One' activity, because I couldn't find."

Conclusions and Implications

- It was determined that newspaper activities affected 4th grade students' attitudes towards science positively.
- According to findings it can be said that newspaper activities balanced the schoolboys' and schoolgirls' attitudes towards science.
- It was seen that newspaper activities affected schoolboys and schoolgirls both positively. But the schoolboys were affected more.
- It was found that students liked newspaper activities and they had fun while doing them. Participating activities voluntary by most of the students is a testament to this. Also Jarman and McClune (2002) determined in their study that teachers use this material for increase interest to the lesson.

This study illustrates the variation of 4^{th} grade students' attitudes towards science after doing newspaper activities. In further studies newspaper activities can be prepared for different grades or different topics and effects on students' science achievements or creative thinking abilities can be investigated.

Teachers can be educated about preparing newspaper activities and using them in science courses. Also preparing a special activity guide for science courses can be useful for science teachers and provide them to use this material more frequently.

References

Bender, M. T. (2006). Resim İş Eğitimi Öğrencilerinde Duygusal Zeka Ve Yaratıcılık İlişkileri. Doktora Tezi, Dokuz Eylül Üniversitesi, İzmir.

Bozdoğan, A. E. (2007). Bilim ve Teknoloji Müzelerinin Fen Öğretimindeki Yeri ve Önemi. Doktora Tezi, Gazi Üniversitesi, Ankara.

Dee-Garrett, S. (2007). Improving Comprehension with Newspapers, 2007 NIE Week Teacher's Guide: Now I Get It!. Retrieved from http://www.mopress.com/_lib/files/07NIEWeek_TeachGuide.pdf

Dee-Garrett, S. (2008). Improving Comprehension with Newspapers, 2008 NIE Week Teacher's Guide: Connect to the World. Retrieved from http://nie.gazettenet.com/NIE%20Week%2008%20teacher%20guide.pdf

Deveci, H. (2005). Sosyal Bilgiler Dersinde Gazete Kullanımı. *The Turkish Online Journal of Educational Technology,* 4:3 Retrieved from http://www.tojet.net/articles/4321.htm

Elliott, P. (2006). Reviewing Newspaper Articles as a Technique for Enhancing the Scientific Literacy of Student-teachers. *International Journal of Science Education*, 28:11, 1245-1265.

Halkia, K., Mantzouridis, D. (2005). Students' Views and Attitudes Towards the Communication Code Used in Press Articles About Science. *International Journal of Science Education,* 27:12, 1395-1411.

Ivan A., & Shibley, Jr. (2003). Using Newspapers to Examine the Nature of Science. *Science & Education,* 12, 601-702.

Jarman, R., & McClune, B. (2002). A Survey of the Use of Newspapers in Science Instruction by Secondary Teachers in Northern Ireland. *International Journal of Science Education,* 24, 997-1020.

Kachan, M. R., Guilbert, S. M., Bisanz, G. L. (2005). Do Teachers Ask Students to Read News in Secondary Science?: Evidence From the Canadian Context. *Science Education,* 90:3, 496-521.

MEB TTKB. (2008) İlköğretim 4-8. Sınıf Öğretim Programlarının Öğrenme-Öğretme Süreçlerinde Gazete Kupürlerinden Yararlanma. Retrieved from http://ttkp.meb.gov.tr

NSF (1998). A Report on the Evaluation of the National Science Foundation's Informal Science Education Program Retrieved from http://www.nsf.gov/pubs/1998/nsf9865/nsf9865.htm

Özçelik, A. (2007). İşbirliğine Dayalı Öğrenmenin Fen Bilgisi Dersinde Başarı, Tutum ve Kalıcı Öğrenmeye Etkisi. Yüksek Lisans Tezi, Osmangazi Üniversitesi, Eskişehir.

Sanderson, P. (2007). Using Newspapers in Classroom. Cambridge: Cambridge University Press.

Sullivan, D. (2002). *Measuring Success! The Positive Impact of Newspaper in Education Programs on Student Achievement.* Newspaper Association of America Foundation. Retrieved from http://nie.yakimablogs.com.php5-2.dfw1-1.websitetestlink.com/wp-content/uploads/pdf/MeasureSuccess.pdf

Vockell, E. L., & Cusick, K. (1995). Teachers' Attitudes Toward Using Newspapers In The Classroom. *Clearing House,* 68, 6.

DEVELOPING SKILLS IN LEARNING AND TEACHING SCIENCE

Josef Trna
Masaryk University

Eva Trnová
Masaryk University

Ivana Vaculová
Masaryk University

Abstract

The paper presents research results of developmental aspects of the students' skills in science education. Science skills are defined, described and classified. Students' science skills get through quite rapid development within school attendance from primary science to upper secondary school science (physics, chemistry, biology, geology). The structure of regulation process within acquiring of science skills is presented. The research objectives are: (1) the measuring of acquiring skills stage during students' skills development in learning science with a view to experimental skills; (2) the analysis of methods of teaching science experimental skills. The comparison of both research outcomes (learning and teaching skills) results in the rules of use of appropriate educational methods which accept developmental aspects. Teachers' science and pedagogical skills play crucial role in the acquiring of students' science skills. These teachers' skills are a part of pedagogical content knowledge (PCK). Research results are presented by concrete science skills and correspondent PCK.

Introduction

Science Skills

Skills are now preferred educational objectives in science education (Ogborn et al., 1996). We need to improve the quality of theory of developing skills and to create innovative teaching methods and tools (Wellington, 1888, 1989).

In pedagogical-psychological literature one can come across several skill-based theories. Comparative studies and other sources (Kraiger, Ford & Salas, 1993; Royer, Cisero & Carlo, 1993) illustrate the existence of four skill-based theories (Švec, 1998, pp. 8-9) which are:

1. Abilities to perform an activity
2. Activity acquirement
3. Internal activity plan
4. Cognitive structure

From a system point of view, which is natural to science education, we lean towards the fourth theory of the skill as a cognitive structure. A skill is therefore defined as a complex ability in the form of a cognitive structure, which is manifested in an observable activity while solving tasks and problem situations.

Skills have a number of characteristics, from which we set out the most important ones:

- *Dependence of skills on learner's gifts*
- *Acquiring of skills by learning and training*

- *Development of skills by spontaneous activities (games etc.)*
- *Close relationship between skill and future activity*
- *Possibility of describing skills due to certain features (correctness and accuracy of solving tasks, etc.)*
- *Possibility to influence the process of acquiring skills by interventionist measures*
- *Influence of auto regulation on acquiring skills*
- *Dependence of an achievable level of acquiring skills on a set of parameters (personality traits of the learner, length of learning, effectiveness of interventionist measures, effectiveness of auto regulation and other)*

In a simplified way we can associate skills with the phrase: "What should students be able to do?". A skill like a disposition cognitive structure is stored in the brain, and is thus observable only indirectly by the subsequent manifested activity.

A skill like a complex cognitive structure has its composition. From a functional perspective, it is appropriate in accordance with (Švec, 1998) to divide a skill into its outer and inner parts. The inner part of a skill is composed of a set of dispositions of the student, in particular:

- dispositions held from birth (abilities),
- learned dispositions (knowledge, habits, experiences),
- styles of cognition, thinking and learning,
- motives, emotions, etc.

These elements of the inner part of a skill form a functional system with a number of relationships. As the core of the inner part of a skill is considered a collection of knowledge concerning a skill, which we call "inner model". This inner model of skill serves in the orientation of a learner when acquiring a skill. For certain skills the student already has developed a preconception form of their inner model.

The outer part of a skill has the form of student performance during the activity concerned and can be identified with an acquired skill of the student (see second skill-based theory). The outer part of a skill, unlike the inner part, is observable and can be diagnosed.

Science Experimental Skills

In everyday life as well as in education we can encounter different kinds of skills. Based on the specific criteria, according to which we analyse individual skills, their different classifications emerge. The main criteria we consider to be:

- Character of prevailing activity
- Range of use
- Complexity of particular activity
- Number of educational subjects, in which they are established
- Degree of creativity

As we are concerned with skills in science, for our future use is the most appropriate the following classification, which brings together all the main criteria for classifying skills (Table 1):

Table 1. Classification of skills.

Criteria	Skills
Character of prevailing activity	rational
	motoric
	communication
Range of use	special
	universal
Complexity of particular activity	simple
	complex
Number of educational subjects, in which they are established	one-discipline
	interdisciplinary
Degree of creativity	reproductive
	productive

In teaching science particularly special skills are acquired (Hodson, 1988). Frequently these are psychomotoric skills, which are closely connected with activities carried out during the exploration process by the scientist. These skills are however modified for education in accordance with educational goals (simplification of performed activities, easier apparatus, etc). As learning in science is based principally on observation and experimentation, we attempted to create an overview of special science experimental skills:

- Designing an experiment
- Designing an experimental apparatus
- Assembling and verification of an experimental apparatus
- Measuring of quantity by a measuring device
- Data collection with the use of tables and construction of graphs
- Subtracting quantity values from graphs and tables
- Data analysis of experimental results
- Carrying out of an experiment

Our research is focused on these science experimental skills. This study presents results of fragmental research development of science experimental skills.

Rationale

Developing of Science Experimental Skills

From an educational perspective the acquiring of skills is of decisive importance. It is thus necessary to discover the role of skills in models of learning. After comparing a number of learning theories (Bransford, 2000), we concluded that it is appropriate to follow the ideas of pedagogical psychologists J. Piaget and P. J. Galperin. The principle shared element of these theories is, simply put, the self-learning action of the pupil during practical activity. That is closely related to his experience and leads to inner thinking operations. As a result, we arrive to the thesis that the acquiring of skills rests principally in activities closely related to experiences.

The student's activity has above all the form of solving tasks and problem situations. These situations typically occur in everyday life (Woolnough & Allsop, 1986). During intervening actions of the teacher during instruction, these situations are artificially connected, and so the teacher's activity is reinforced. A higher level signifies auto regulated actions when the learner knowingly solves tasks and problems, during which he/she is acquiring the relevant skills.

The process of skills development is also influenced by a number of additional personal factors, such as: self-reflection, anticipation, motivation, will, temperament etc.

In accordance with (Talyzinova, 1988) in the process of developing a skill, five stages of development can be identified:

(1) **M**otivational stage**:** The student's motivation when acquiring skills is the basic prerequisite, in the same way as during the entire learning process. As a skill is closely related to the activity, the teacher has at his disposition more motivational techniques than during acquiring of knowledge. A very important motivation is practical utilisation of skills when solving task and problem situations, especially in everyday life.

(2) **O**rientation stage**:** The student must acquire the necessary knowledge as the base for the intellectual component of the skill. Rational and motoric components of a skill are usually directed with the help of instructional methods. What is meant is that word (verbal or written) information is connected with image information. Precisely, practice of necessary routines is concerned. The interconnection of the sensory and motoric components can have a motivational effect and ads versatility and durability to the skill. In the next phase the individual's thought process is more effective connected with manipulation and in the final phase then the individual's thought process connected with thought manipulation (imagining realisation of the manipulation).

(3) **S**tabilisation stage**:** Here the solving of simple applied tasks is concerned, during which the skill being acquired is used. These tasks are solved by the student according to written or verbal instructions (for example laboratory instructions).

(4) **C**ompletion stage**:** Student should on the basis of acquired orientation stage (functioning as auto-regulation) be able to solve more complicated tasks by using the acquired skill and by using previously acquired knowledge, routines and skills. In the older age development stage of a student, one can take advantage of a thought simulation experiment, which the student is aware of. Here, self-diagnosis and inner motivation is of great importance.

(5) **I**ntegration stage**:** Including this skill in the entire complex of skills takes place above all when solving interdisciplinary problem tasks and projects.

Research Questions

It is necessary to consider the above mentioned stages of acquiring a skill when planning and realizing classroom instruction. This is why it is necessary to examine the development of science skills of students. The practical application of this research should be teaching methods and strategies. Empirical research can demonstrate real life situations of teaching and learning of science skills at secondary schools in the Czech Republic.

The research objectives and research questions are:

(i) **M**easuring of acquiring stage of skill during students' developing skill in learning science (at the age of 8 -19) with a view to science experimental skills:

Does the stage of acquired skill depend on the age of student? Which age of student corresponds with the concrete stage of developing skill?

(ii) **A**nalysis of methods of teaching skills:

Do science teachers apply teaching methods containing the stages of developing skill?

(iii) **C**omparison of both (i) and (ii) research outcomes (learning and teaching parts of skill acquiring) results to the rules of use appropriate educational methods accepted developmental aspects of skill:

Do science teachers use appropriate teaching methods and tools for effective skill development?

The purpose of combination of these three empirical questions is to result in primary information concerning the development of science experimental skills of students and finding out whether science teachers react to this skill development.

Methods

Sufficient methods for our research objectives were:

(I) **W**orksheets for students (students solve the set of tasks with different skill stage)
(II) **V**ideo study of lectures (analysis of video recorded lectures with teaching skills)
(III) **C**omparison of worksheets and video study results

Individual fragmented research methods were applied in the following way (Trna & Trnova, 2008):

Worksheets for Students

Students were given activities based on worksheets. These worksheets were composed of tasks, by which individual phases of acquiring experimental skills were realized (stabilisation, completing and integration). We did not include in worksheets the first two stages of acquiring skills (motivation and orientation), because as elementary stages of acquiring skills they should be mastered by the student. Tasks corresponding to individual stages being investigated can be described as follows:

(A) **S**imple reproductive task (stabilisation stage): Student progressed exactly according to the instructions and there was no need to use his creative thinking.
(B) **M**ore creative task (completing stage): Student should use his creative thinking not only according to the detailed description.
(C) **C**omplex problem task or project (integration stage): The skill had to be creatively integrated into a structure of other skills and used as a complex skill.

Empirical research question (i) and verification of the hypothesis concerned about the dependence of stage achievement of acquirement of students skills on the age was implemented by the different levels of tasks leading to acquirement of the temperature measuring skill (Trna & Trnova, 2006):

Students at the age of 8 – 11:
(A) Tasks (stabilisation stage):
Students solve tasks on the basis of written and verbal instructions. They measure the human body temperature and air temperature in the classroom.
(B) Tasks (completing stage):
Students compare temperatures by measuring temperatures inside the school, outside the school building (at different points of the compass) and at different levels above and below ground in the school garden.
(C) Tasks (integral stage):
A project on temperature measuring was oriented to finding the relationship between air temperature and the behaviour of animals (reproduction) or plants (generation of fruits).

Students at the age of 12 – 15:
(A) Tasks (stabilisation stage):
Students solve tasks on the basis of written or verbal instructions (laboratory directions). The stabilisation stage should include, in addition to temperature measuring at one moment in time, also repeated temperature measurement based on time (air temperature measuring during the day). Formation and evaluation of graphic dependence (interval of temperature increase and decrease, maximal and minimal temperature) can be included as significant skills. Application of maximal-minimal thermometers as a supplement is recommended.
(B) Tasks (completing stage):

Students solve experimental tasks involving temperature from a range of natural substances and phenomena. We included tasks in the form of the following laboratory experiments: Physics: water melting and boiling temperature curve. Chemistry: changes of steam temperature during liquid distillation. Geography: meteorological measurements.
(C) Tasks (integral stage):
Only gifted students were able, sometimes, to solve a creative project task (e.g. project Globe).

Students at the age of 16 – 19:
(A) Tasks (stabilisation stage):
Students solve tasks on the basis of written laboratory directions. They measure temperature on the screen of the school weather station; they check the water temperature in the school aquarium etc.
(B) Tasks (completing stage):
Students, after achieving the stabilisation stage (which takes effect of self regulation), can manage complicated problem solving by means of their acquired skills, as well as their conceptual knowledge. We can now use more complicated temperature measuring tasks: Physics: Thermometers calibration. Chemistry: Impact of temperature against a chemical reaction process. Biology: Contraceptive method based on a basal body temperature measuring. Geography: Thermal changes in a soil profile.
(C) Tasks (integral stage):
A project with an interdisciplinary character is an appropriate accession to the final stage of the skill development. This skill is added complementarily to all skills gained by students by the solving of interdisciplinary problems. The temperature measuring skill can be integrated, for example, into environmental problem solving (global warming, etc.).

Video Study of Lectures

Verification of the hypothesis of the concerned empirical research question (ii) in relation to the application of the stages of development of science skills was based on a video study.

We used observation based on the video recording of physics lectures with the following analysis of the recording, i.e. video study (Vaculova & Trna, 2008). It concerns 27 lessons on the topic of composition of forces taught by 8 teachers in 8 classes (with 177 students in total) at lower secondary schools (age of students 13-14). All the teachers were qualified to teach physics and had experience from 2 to 28 years. The video recordings of every teacher were made in 2-4 successive lessons. The observation of the video recording was structured in the categories which had been specified before the observation started (see Table 2).

Table 2. Categories and coding.

Category	Numeric codes and subcategories
Phase of the process of skill acquirement	1 orientation phase 2 stabilisation phase 3 completion phase 4 integration phase

The video recordings were coded and quantified in Videograph programme (Rimmele, 2002) at ten-second interval.

Results

Recognized percentage of acquiring skill stage bear evidence of the dependence on the age of student (see Table 3).

Table3. Measuring of acquiring stage of skill during skills development in learning science.

Age	*Stabilisation stage (%)*	*Completing stage (%)*	*Integration stage (%)*
8 – 11	84	10	0
12 – 15	91	43	12
16 – 19	97	66	19

Analysis of tasks, which was given by the teachers to the students, resulted in the finding that development stages of acquiring skills are not appropriately developed (Figure 1 and 2).

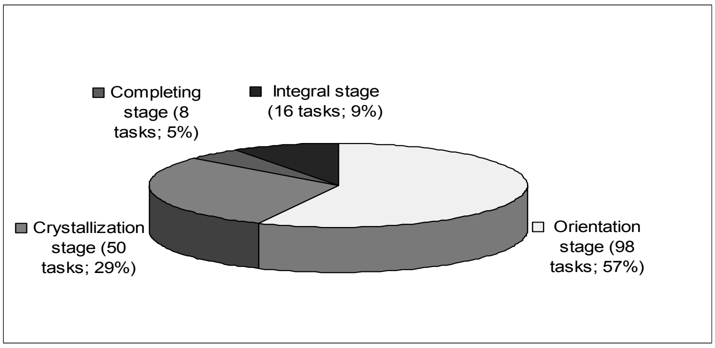


Figure 1. Total frequencies of the tasks according to the type of the phase of skills acquirement.

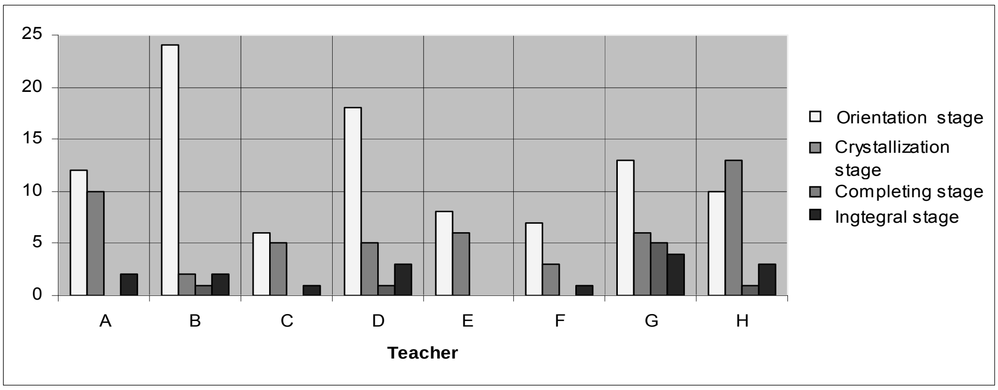


Figure 2. The tasks frequencies according to the type of the phase of skills acquirement by the particular teachers.

When comparing the results of the research study of the development of experimental skills of students (i) with methods of their acquiring, which are used by science teachers (ii), a contradictions is found. Teachers are either unfamiliar with the stages of developing skills or they often ignore them when teaching.

Conclusions and Implications

Our research outcomes result in the rules of use appropriate educational methods which accept developmental aspects. Science teachers do not use sufficient teaching methods for effective learning science skills of students (with a view to experimental science skills). They do not have good pre-service and/or in-service preparation for developing skills in learning and teaching science.

If we want to change this rather bad situation, it is necessary to accept the existence of development stages of skills. Teachers have to accept and use in teaching/learning science rules of developing skills:

Students at the age of 8 – 11: Students were capable of visual operation with concrete things. Therefore it is possible, at best, to reach to the stabilisation stage of a new skill. The motivational stage was found to be the critical role. The orientation stage could not be too time-consuming, or the content too difficult.

Students at the age of 12 – 15: Theses students can be involved in abstract thinking. We can reach even to the completing stage. The motivational stage plays still an important role. The orientation stage can be more time-consuming as well as encompassing more difficult content and abstract concepts can be acquired and structured. Students' manual handling skills are at a high level. The core of skills acquired is at the stabilisation stage.

Students at the age of 16 – 19: Student achieves a high level of personal development. Thus the completing stage involving problem tasks solving tasks can be mostly realised. The tasks can be in the form of project with strong interdisciplinary character. It is important not to forget repetition and completion of all previous stages within the undertaking of the project teaching.

Our research provided evidence that science skills (especially experimental skills) are an important subject for research and development of appropriate teaching/learning methods and tools in science education.

Acknowledgements

The paper was created and supported within the project projects "Special Needs of Pupils in Context with Framework Educational Program for Primary Education"(MSM0021622443).

References

Bransford, J. D. (Ed) (2000). *How People Learn. Brain, Mind, Experience, and School (Expanded edition).* Washington, DC, USA: National Academies Press.

Hodson, D. (1988). Experiments in science and science teaching. *Educational Philosophy and Theory,* 20, 253– 266.

Kraiger, K., Ford, J. K & Salas, E. (1993). Application of Cognitive, Skill-Based, and Affective Theories of Learning Outcomes to New Methods of Training Evaluation. *Journal of Applied Psychology*, 78(2), 311-328.

Ogborn, J., Kress, G., Martins, I. & McGillicuddy, K. (1996). *Explaining Science in the Classroom.* Milton Keynes: Open University Press.

Rimmele, R. (2002). *Videograph. Multimedia-Player zur Kodierung von Videos.* Kiel: IPN.

Royer, J. M., Cisero, Ch. A. & Carlo, M. S. (1993). Techniques and procedures for Assessing Cognitive Skills. *Review of Educational Research*, 63(2), 201-243.

Švec, V. (1998). *Klíčové dovednosti ve vyučování a výcviku.* Brno: Masarykova univerzita.

Talyzynova, N. F. (1988). *Utváření poznávacích činností žáků.* Praha: SPN.

Trna, J. & Trnova, E. (2008). Diagnostics of the Analytical Observation and Preconceptions in Science Education Particularly Focused to Gifted Students. In Proceedings of XIII. IOSTE Symposium the Use of Science and Technology for Peace and Sustainable Development*, (pp. 1240-1246).* Izmir: Buca Faculty of Education.

Trna, J. & Trnova, E. (2006). Science Experimental Skills under Development. In Europe Needs More Scientists. The Role of Eastern and Central European Science Educators *(pp. 77-84).* Tartu: University of Tartu.

Vaculova, I. & Trna, J. (2008). Video-study of Physics Education: the Role of Teaching Tasks in the Process of Skills Acquirement. In Research in Didactics of the Science*, (pp. 182-185***).** Krakow: Uniwersytet pedagogiczny im. Komisji Edukacji Narodowej.

Wellington, J. J. (1988). The place of process in physics education, *Physics Education,* 23, 150–155.

Wellington, J. J. (ed.) (1989). *Skills and Processes in Science Education.* London: Routledge.

Woolnough, B. & Allsop, T. (1986). *Practical Work in Science.* Cambridge: Cambridge University Press.

TEACHING SIMPLE MACHINES TO COLLEGE STUDENTS THROUGH LEGO™ ENGINEERING DESIGN CHALLENGES

Ismail Marulcu
Boston College

Michael Barnett
Boston College

Abstract

Despite the lack of design-based curriculum for elementary school students and teachers' low self-efficacy in regard in teaching design, there appears to be a strong demand for engineering (design) education in schools. We believe that deepening science learning and increasing integration of engineering education in elementary levels can be accomplished concurrently with using engineering based curricular modules including engineering design challenge practices in elementary science classrooms. In this study, we test the hypothesis that engineering-design based curriculum modules can deepen science learning by helping students to achieve higher order knowledge. In this phase of the study, we aim to investigate college students' (especially pre-service teachers') preconceptions and understanding of simple machines. Hence, later on, this will give us an opportunity to compare elementary students' preconceptions with college students' preconceptions on simple machines. The fifth grade LEGO engineering-based simple machines module was modified to college level, and it was piloted in a college located in a big city in the Northeastern region of the US for two weeks. We found that students who participated in the study show significant growth in content understanding, and we identified several possible misconceptions of college students on simple machines.

Introduction

Understanding and using technology and technological tools that were made possible by scientific understanding and engineering design is getting more and more important in today's world, which is competitive in terms of skills for using scientific and technological knowledge.

Increasing technology oriented daily life practices expands the importance of science and technology education starting with the elementary grades (NRC, 2006). Standardized test results mirror the need for deeper understanding in elementary science education. In the 2005 National Assessment for Educational Process (NAEP) test, 73% of fourth graders in the U.S. scored under proficient level in science (NCES, 2005). This is not surprising since Anderson and Mitchener (1994) showed that elementary teachers do not have a strong content knowledge in science, [especially in physical sciences] and Cochran and Jones (1998) showed that their confidence level of teaching science is pretty low.

Not only improvements in science education, but also integration of engineering education are needed in elementary classrooms. The American Society for Engineering Education (ASEE) conducted a survey on K-12 teachers examining the need for engineering education in schools (Douglas et al., 2004). 90.3 % of teachers agreed that 'understanding more about engineering can help them become a better teacher', and 84 % agreed that 'their students would be interested in learning engineering'. Despite this strong demand for engineering education in K-12 schools, students lack of engineering awareness. A recent study reported that less than 15% of 504 elementary students in one Massachusetts district, correctly identified 'creating ways to clean water' as something that engineers do, while over 70% incorrectly chose 'drive machines' as a special engineering task (Cunningham, Lachapelle, & Lindgren-Streicher, 2005).

Various studies showed that many students do not enjoy science classes because they find them uninteresting and irrelevant (e.g. Williams, Stanisstreet, Spall, Boyes, & Dickson, 2003, Woolnough, 1994, and Briggs, 1976). However, LEGO use is a very powerful way of adding fun and motivating students in science classrooms. It can potentially motivate teachers by helping them increasing students' achievement. This study is unique in its integration of LEGO engineering tools and triarchic teaching and assessment methods (Sternberg, 1985). In this phase of the study, we aim to investigate college students' (especially pre-service teachers') preconceptions and understanding of simple machines. Hence, later on, this will give us an opportunity to compare elementary students' preconceptions with college students' preconceptions on simple machines. The driving questions for our work have been:

1. Is engineering design based curricular module on simple machines successful in teaching pre-service teachers?
2. What are the misconceptions of pre-service teachers on Simple Machines?
3. Does LEGO-engineering design based instruction help pre-service teachers to overcome their misconceptions?

Background

National standards and benchmarks for science education encourage students to get familiar with engineering design by engaging in design activities. For example, the National Science Education Standards (NSES) argues that "children's abilities in technological problem solving can be developed by firsthand experience in tackling tasks with a technological purpose" (NRC, 1996 p. 135). Similarly, under Project 2061, the American Association for the Advancement of Science (AAAS) argues that

Perhaps the best way to become familiar with the nature of engineering and design is to do some. By participating in such activities, students should learn how to analyze situations and gather relevant information, define problems, generate and evaluate creative ideas, develop their ideas into tangible solutions, and assess and improve their solutions. To become good problem solvers, students need to develop drawing and modeling skills, along with the ability to record their analyses, suggestions, and results in clear language (AAAS, 1993, p. 48).

Our engineering-design based modules are designed in to engage students in an engineering-design process with the end goal of solving real life problems while engaging students with basic scientific content. The basic blueprint of the units is based upon the Learning by Design™ model developed by Kolodner and colleagues at Georgia Technical University (Hmelo, Holton, & Kolodner, 2000; Kolodner, Camp, Crismond, Fasse, Gray, Holbrook, Puntambekar, Ryan, 2003; Kolodner, Gray, & Fasse, 2003). This model is based on five principles, which emphasize the daily practices of skilled practitioners such as scientists, engineers, and industrial designers. These principles include foregrounding of skills and practices, practicing, establishing need, making recognition of the need to use procedures automatic, and establishing and enforcing expectations (Kolodner, 2002). The design challenges generates a motivation for learning the science content, and engaging in design challenges provides natural settings and procedures for practicing science and design skills. The need for creating a working model of design ideas helps students to identify their imperfect and poor conceptions and to restore them while natural iterations in the design processes allow students to apply and test new conceptions. Also, the collaborative practices of team work provide students the opportunities to improve in communicating ideas and results.

Engineering design activities provide students not only an environment for learning science content but also opportunities for application of the knowledge that they gain in different situations and for engaging in the practices of scientists (Kolodner, 2002) and engineers. In Learning By Design (LBD), the design challenge provides a reason for learning the science content, and engaging in the challenge provides a natural and meaningful venue for using both science and design skills. The need to make one's design ideas work provides opportunities and reasons for students to identify incomplete and poor conceptions of science content and to debug those conceptions; the iterative nature of design provides opportunities to apply and test new conceptions; and the collaborative nature of design provides opportunities for team work and the need to communicate ideas and results well (Kolodner, 2002, p. 3).

Kolodner found that students who learned science with LBD learned content as well or better than the students who learned the content with traditional methods. Along with content knowledge, LBD students have learned many skills that scientists and designers often use in their profession. In another study, Kolodner, Gray, and Fasse (2003) conclude that LBD students involved in collaboration, communication, informed decision making, and design of investigations more skillfully than the students in control groups.

Design activities have been implemented in elementary school settings (Penner et al., 1998; Roth et al., 2001), in out of school settings (Davis et al. 1997), in design competitions (Davis & Masten, 1996; Padgett, 1997), in high school settings (Barnett, 2005), and in undergraduate engineering programs (Cross, 2006; Ringwood et al., 2005). To date, however, little research has been exploring whether design based activities are effective in pre-service teachers' learning science.

To adopt new instructional methods effectively in science teaching, those methods should be introduced to teachers in their pre-service education (Bers & Postmore, 2005). With this inspiration, Bers and Postmore introduced a partnership model for pre-service early childhood teachers to learn how to develop, implement and evaluate curricula in mathematics, science and technology. They used Lego Robotics materials as an instructional tool. In this model, early childhood teachers were teamed up with engineering students in three different forms of partnership models: the developer's model, the external consultant's model, and the collaborator's model. Among these partnership models, the most successful but the most time consuming one was the collaborator's model. Bers and Postmore conclude that the partnership models allowed pre-service teachers to realize how technology could potentially improve their educational practices and what skills they need to use technology effectively.

There are only a small number of studies that investigate students' or teachers' understanding of simple machines in the science education literature. In our search of the research base, we found only one study that examined pre-service teacher knowledge of simple machines. Joseph Taylor (2001) investigated an instructional sequence that encourages conceptual change in bicycle science for pre-service elementary teachers. His initial assessment showed that pre-service elementary teachers' understanding of some concept such as work and torque is problematic. His instructional sequence includes making ideas explicit, brief discussion of essential concepts and terminology, investigation of essential relationships, and integration and application: a bicycle (drive train) investigation. Of twenty-nine pre-service elementary teachers participating in the study, twenty-three of them correctly identified that simple machines trade input distance for input force or vice versa. Taylor concluded that his instructional sequence encouraged conceptual change. A parallel procedure to inquiry science and its inclusion of design and modeling activities as well as investigation of the real object (bicycle) seems the reasons for success of Taylor's instructional sequence in promoting conceptual change.

In summary, we are building off the extensive work of Kolodner (2002) who identified three challenges to success in teaching by engineering design as teacher preparation, assessment of skills (learning), and time. In particular, Kolodner noted that engineering design is a new and important method for teachers as well as their students and in order to be successful teachers need the experience of learning by engineering design to teach with engineering design. For this reason, we have investigated whether engineering design is an effective instructional tool to teach science to pre-service teachers.

Methods

A mixed-methods approach is going to be used to answer the research questions. We see numerous advantages of using a mixed-method approach. These include having different perspectives with each method, being able to gather more in-depth information, and being able to ensure reliability. Lincoln and Guba (1986) suggested triangulation as one of the ways of increasing the reliability of researcher interpretations. We triangulated data through multiple sources, including direct observations, tests, interviews, and student laboratory reports.

Study Context

During the 2007-2008 academic year, we implemented a portion of the Transforming Elementary Science Education through Lego Engineering Design (TESLED) work in one of our researcher's 'The Living Earth II' course at a college in a big northeastern city. This course is one of the most selected courses as a required science course for college students and is specifically designed for future elementary science teachers. Among 130 students taking the course, pre-service teachers constituted the majority and students from other colleges were the minorities. An adapted and brief version of the LEGOTM engineering-based simple machines module, which was developed by our research team, was implemented in 2 weeks of the course laboratory session, which occurs once a week. During the first week of the study, students explored 5 types of simple machines that are levers, pulleys, wheel and axles, inclined planes, and gears by using engineering design based activities. Then, students, in groups were asked to solve three engineering design challenges on simple machines during the second week of the study.

Data Collection and Analysis

We collected data from multiple sources. We prepared and administered pre-post surveys by using the Survey Monkey, an online tool for creating web surveys, to measure pre-service teachers' understanding of simple machines. The pre- and the post-test included 10 multiple choice and 5 open-ended questions, which are the same. Also, the post-test included one more open-ended question that asks students to express what they learned after the instruction. We administered pre-post interviews to explore pre-service teachers' understanding of simple machines in depth as well. We observed and video-taped a sample of laboratory sessions. Lastly, we collected students' laboratory reports. These laboratory reports included pre-service teachers' reflections to the challenge questions in each simple machine activity and pictures of their final challenge engineering designs.

We analyzed pre and post surveys by using SPSS, and we used an interpretive framework to analyze our data. We transferred the statistical survey results provided by the Survey Monkey to excel worksheets and draw their graphs after eliminating double and triple entries. Also, in terms of the surveys we entered the data into SPSS version 11.0 and conducted pre-post paired t-tests to evaluate the impact of the engineering-based design challenges experience. Two of the researchers in our research team scored the open ended items in the pre- and post-surveys and the interview items. The items that were scored differently by different researchers were scored again together and consensus is reached.

In terms of analyzing the student interactions, we watched the video tapes and broke the student discussion down into interactive episodes. We defined an episode to be a conversation in which the same idea, topic, or concept was being discussed. These video episodes were then grouped and analyzed holistically with the goal of identification of particular tensions or contradictions that either supported or inhibited learning.

Results

Open-ended questions and interviews were scored by two researchers, and Cronbach's Alpha for open-ended questions was found 0.7236, and for interviews was found 0.9133.From the analysis of test results, we found that college students, who participated in the study, showed significant growth in their test scores ($t = 3.36$, $p < 0.001$ for multiple choice items on the test and $t = 2.47$ for the open-ended items on the test).

Table 1 shows pre-service teachers' mean scores on pre- and post-tests and interviews and gains between them. Pre-service teachers significantly improved their scores in both multiple choice and open-ended sections in the post-test and in the post-interview. This means that there was a significant improvement in pre-services teachers' content understanding of simple machines.

Table 1. Pre-service teachers' mean scores on pre- and post-tests and interviews and gains between them.

	Multiple Choice Items (N=85)			**Open Ended Items (N=85)**			**Interview Items (N=10)**		
	Pre Score	Post Score	Change	Pre Score	Post Score	Change	Pre Score	Post Score	Change
Average Score	6.01	6.73	0.72	5.18	6.13	0.95	7.5	12.6	5.1
Standard Deviation	1.401	1.577		2.274	2.935		1.434	1.838	
Significance			t= -3.363 p<0.001			t=2.47 p<0.016			t= p<0.000

In the pre- and post-interviews the first question we asked required defining simple machines and selecting them among different objects. Follow-up questions included why they picked a particular object as a simple machine and what the purposes of simple machines are. Seventeen college students –most of them were pre-service teachers- were interviewed. Lego-engineering activities seemed to maintain conceptual development as the students increased their average score from M = 1.117 (SD=0.332) to M = 2.352 (SD=0.606).

Pre-Service Teachers' Misconceptions about Simple Machines.

Through our analysis of tests and interviews results, we identified three possible misconceptions that could cut across the both study contexts. These misconceptions are that simple machines change the amount of work need to do a task, bigger pulley requires less force than a smaller pulley to pull things, and the angle we pull the rope in a pulley system changes the amount of force required to pull the load.

Simple Machines Change the Amount of Work You Do.

Subjects who hold on to this idea expressed their thinking in different ways. Some of them directly stated that simple machines reduce the amount of work needed to do a task while others stated that simple machines reduce the energy we use to do a task. The percentage of pre-service teachers who had this misconception decreased significantly in the post-test. Figure 1 shows the distribution of answers to the related question in the pre- and the post-test. In the pre-test, 71 percent of the subjects expressed that a simple machine decreases the amount of work that you need to do to move an object, while this decreased to 47 percent in the post-test. Correspondingly, the percentage of subjects who picked the correct answer (a simple machine changes the amount of force that you need to apply to move an object) increased from 26 percent to 50 percent. This shows that, even though Lego-engineering design challenges could not correct all the subjects' misconceptions, it could successfully transform a substantial portion of their misconceptions to correct conceptions.

Bigger pulley requires less force than a smaller pulley to pull thing.

It seems that this idea stayed solid throughout the instruction. Figure 2 shows the distribution of answers to the related question in the pre- and the post-test. In the pre-test 24 percent of the subjects expressed this idea while this percentage fell only one point in the post-test. Percentage of the subjects who correctly chose the idea: size of the pulley does not any difference increased from 43 percent to 58 percent after the instruction. However, it seems that the subjects who changed their ideas in the post-test were the ones who thought that the weight will move slower or faster as we pull. The reason why the subjects held on to the idea that a bigger pulley requires less force than a smaller pulley might be that the subjects see simple machines as objects that help humans do tasks easier. Therefore, they might think that once we use a bigger simple machine, we use less force.

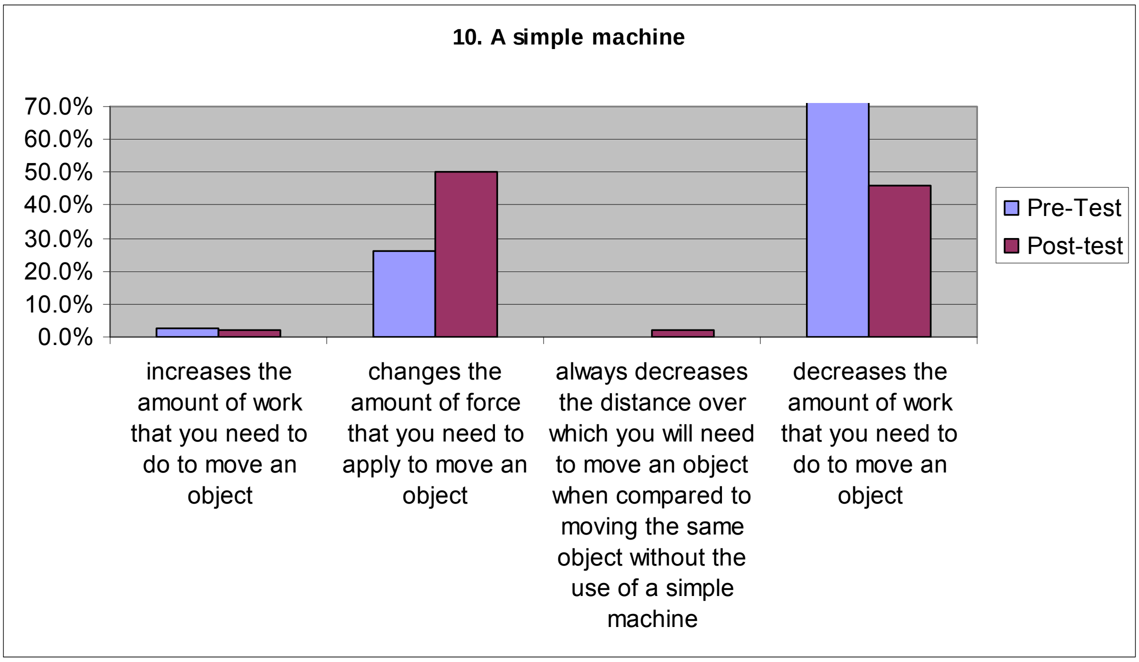


Figure 1. The distribution of answers to question 10 in the pre- and the post-test.

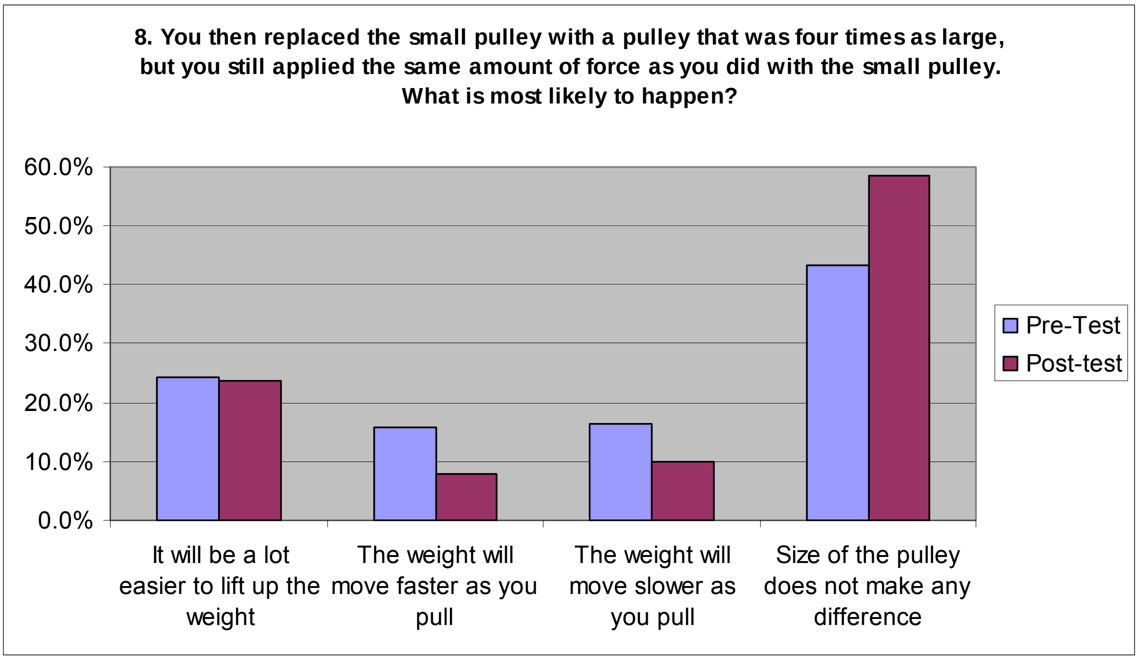


Figure 2. The distribution of answers to question 8 in the pre- and the post-test.

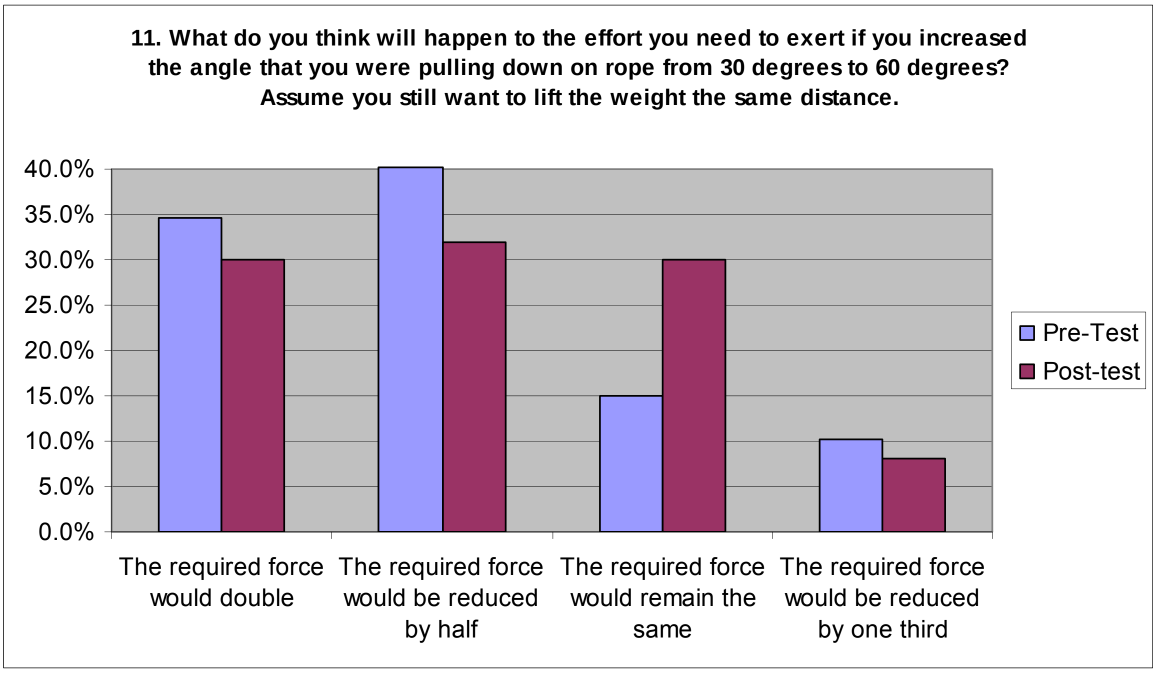


Figure 3. The distribution of answers to question 11 in the pre- and the post-test.

Sample Conversations about Concepts.

It is seen from the test results that many pre-service teachers have a misconception that simple machines decrease the work we do. In the pre-interviews, seven out of seventeen subjects used expressions reflecting the misconception, while only two of them and two others used expressions reflecting the misconception in the post interviews. For example, before the instruction, Student 1 believed that simple machines help us "do more work with like less energy".

Student 1 Pre-Interview:

Researcher: So why do you think scissors is a simple machine?

Student 1: I know that the definition of a machine is something that helps do more work with like less energy. I think it is easier to cut with scissors than to try with hand.

In the post-test, Student 1 expressed that simple machines do not change work.

Student 1 Post-Interview:

Researcher: What do simple machines do?

Student 1: Ahhm.. They either allow you to change the direction of a force or to reduce the amount of force needed to do an amount of work.

Researcher: Do they reduce the amount of work?

Student 1: Ahh.. no.. the work is the same.

It is evidenced that Lego-engineering design based instruction helped him overcome the misconception. Another case for the reconstruction of the first misconception was Student 2's case. The following sequence shows that during the pre-interview, Student 2 believed that simple machines reduce work by doing some of it for us, *Student 2 Pre-Interview:*

Researcher: So imagine that there is an object here (pointing the load side of a one fixed pulley system) and if you pull…

Student 2: Yeah, it would go up higher.

Researcher: So, in that case do you use less force than the weight of the object?

Student 2: I think, technically you are using the same amount of force because the object itself doesn't have less mass of anything like that. So it requires the same amount of force. But it is the amount of energy that is less because you have the simple machine doing some of the work for you.

However, during the post-interview, she expressed that simple machines do not reduce work and they just reduce the force you exert.

Student 2 Post-Interview:

Researcher: What do simple machines do?

Student 2: They don't make work less. They just make work easier to do, so technically if you did the formula to find out how much work you use whether to use a simple machine or not.

Researcher: Can you describe work? What is work?

Student 2: Work is force times distance. So it is the amount of force you have to apply on an objects multiplied by the distance you have to move the object. And then, multiply that you get the work. So it is… Yes, work is force times distance. So using simple machines does not make work more or less. It just helps you do the work. So like if you are trying to push a box it is on the ground or something. If you push it just when it is flat on the ground, it is going to be harder to do than if you had on like... instead of like a wheel and axle or something like that because it.. The simple machines help you to move it, but you are still putting the same amount of force into what you have to do. So that the numerical value of work is the same. It just feels easier to.

Researcher: So what changes?

Student 2: So if you have like, an inclined plane. The force maybe... If you lift it up… Like, straight up to go… Say 10 meters higher, the force might be say I don't know 5 Newtons. And then the distance is 10 meters. The work would be 50 joules. Then if you used an inclined plane the force could be only 2 Newtons, but the distance because you are using an inclined plane increases, so since the distance increases the work stays the same. So the distance could be like 25 meters and then you still get the same; 50

Conclusions and Implications

It is possible and important to integrate engineering design education into pre-service teaching (Kolodner, 2002, Bers & Postmore, 2005). Our curriculum materials support students in learning content along with engineering design (Kolodner, 2002). Learning science content through engineering design tasks appears challenging but given limited amount time such an approach does appear to have significant potential. In engineering design activities, students learn both engineering design and science content. It is important that support is given to the student actions in their own construction of knowledge as they emerge from this activity / learning systems. This activity emphasizes the "situated nature of cognition and meaning (Barab et al, 2002, Barab & Plucker, 2002)" suggesting a "reformulation of learning in which practice is not conceived of as independent of learning and in which meaning is not conceived of as separate from practices and contexts in which they are developed (Barab et al, 2002, Lave & Wenger, 1991)."

Gaining a conceptual understanding of simple machines is important for pre-service elementary students because simple machines is a standard curriculum unit taught in many school systems (i.e. STC, FOSS). We noticed that a number of concepts (e.g. work, torque, conservation of energy) around simple machines are difficult for pre-service teachers to understand. Therefore, these concepts should be provided extra attention by teacher educators and instructional designers. Although, not all the participants in the study were pre-service teachers, it suggests that engineering-design based curricular modules have a potential for deepening science learning as well as increasing integration of engineering education in elementary levels. It also suggests that engineering-design based curricular modules could be effective in eliminating pre-service teachers' misconceptions and transforming them into scientifically correct conceptions.

As a result of our findings, we are undertaking additional design work is required to develop focused and effective activities that help ameliorate pre-service teachers' misconceptions around simple machines so that they can teach their future students.

Acknowledgements

This work is supported in part through a National Science Foundation (NSF) Research and Evaluation on Education in Science and Engineering (REESE) program (Grant # 0633952). The authors would like to thank Kristen Wendell, Kathleen Connolly, Chris Wright, Linda Jarvin and Chris Rogers from Tufts University for their valuable contributions.

References

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. New York: Oxford University Press.

Anderson, R. D., & Mitchener, C. P. (1994). Research on science teacher education. In D. L. Gabel (Ed.), *Handbook of Research on Science Teaching and Learning*. New York: Macmillan.

Barab, S.A., Barnett, M., Yamagata-Lynch, L., Squire, K. Keating, T. (2002). Using activity theory to understand the systemic tensions characterizing a technology-rich introductory astronomy course. Mind, Culture, and Activity, 9(2), 76-107.

Barab, S. A., & Plucker, J. A. (2002). Smart people or smart contexts? Cognition, ability, and talent development in an age of situated approaches to knowing and learning. Educational Psychologist, 37(3), 165–182.

Barnett, M. (2005). Engaging inner city students in learning through designing remote operated vehicles. *Journal of Science Education & Technology*, *14*(1), 87-100.

Bers M. U., & Portsmore, M. (2005). Teaching partnerships: Early childhood and engineering student teaching math and science through robotics. *Journal of Science Education and Technology., 14*(1), 59–73.

Briggs, B. H. (1976). Student attitudes to physics. [Electronic version*]. Physics Education, 11,* 483-487.

Cochran, K., & Jones, L. (1998). The subject matter knowledge of preservice science teachers. In B. J. Fraser & K. G. Tobin (Eds.), *International handbook of science education* (pp. 707-718). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Cunningham, C.M., Lachapelle, C. & Lindgren-Streicher, A. (2005). "Assessing elementary school students' conceptions of engineering and technology." Proceedings of the ASEE Annual Conference.

Cross, N. (2000), *Engineering Design Methods – Strategies for Product Design*, Chichester : John Wiley & Sons,.

Davis, M.L. and Masten, S.J., (1996) Design competitions: does multidisciplinary contribute to the team building experience? in *Proceedings of the 26th Frontiers in Education Conference*, *I* (pp. 276–279).

Douglas, J., Iversen, E., & Kalyandurg, C. (2004). Engineering in the K-12 classroom: An analysis of current practices & guidelines for the future. Washington, DC: American Society for Engineering Education.

Hmelo, C.E., Holton, D.L., & Kolodner, J.L. (2000). Designing to learn about complex systems. *Journal of the Learning Sciences*, 9(3), 247 - 298.

Kolodner, J., Camp, P., Crismond, D., Fasse, B., Gray, J., Holbrook, J., et al. (2003). Problem-Based Learning Meets Case-Based Reasoning in the Middle-School Science Classroom: Putting Learning by Design(tm) Into Practice. *Journal of the Learning Sciences*, *12*(4), 495. http://search.ebscohost.com, doi:NO_DOI

Kolodner , J. L., Gray, J. T., & Fasse, B. B. (2003). Promoting transfer through case-based reasoning: Rituals and practices in Learning by Design™ classrooms. *Cognitive Science Quarterly* 3(2), 119-170.

Kolodner, J. (2002). Facilitating the learning of design practices: Lessons learned from an inquiry into science education. *Journal of Industrial Teacher Education*, *39*(3), 9-40. Retrieved April 1, 2009, from Education Research Complete database.

Lave, J., & Wenger, E. (1991). Situated learning: Legitimate peripheral participation. New York: Cambridge University Press.

Lehrer, R., & Romberg, T. (1992). Exploring children's data modeling. Cognition and Instruction, 14(1), 69-108.

Lincoln, Y. S. and Guba, E. G. (1986) But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. New Directions for Program Evaluation, 30, 73–84.

National Center for Education Statistics (2006). The Nation's Report Card: Science 2005 (NCES 2006-466). U.S. Department of Education. Washington, DC: National Center for Education Statistics.

National Research Council [NRC] (1996). *National Science Education Standards*. Washington, DC: National Academy Press.

Padgett, W.T. (1997). Teaching design through design competition, in *Proceedings of the 27th Frontiers in Education Conference*, *III* (pp. 1477–1480).

Penner, D. E., Lehrer, R., and Schauble, L. (1998). From physical models to biomechanics: A design-based modeling appraoch. *The Journal of the Learning Sciences* 7: 429–449.

Ringwood, J., Monaghan, K., & Maloco, J. (2005, March). Teaching engineering design through Lego ® Mindstorms™.*European Journal of Engineering Education*, *30*(1), 91-104. Retrieved February 16, 2008, doi:10.1080/03043790310001658587

Roth, W.-M., Tobin, K., & Ritchie, S. (2001). *Re/ConstructingElementary Science*, Peter Lang, New York.

Taylor, J. (2001). Using a practical context to encourage conceptual change: An instructional sequence in bicycle science. *School Science & Mathematics*, *101*(3), 117. Retrieved February 16, 2008, from Education Research Complete database.

Williams, C., Stanisstreet, M., Spall, K., Boyes, E. and Dickson, D. (2003). 'Why aren't secondary students interested in physics?', *Physics Education*, 38(4), pp. 324 329

Woolnough, B. E. (1994). Why students choose physics, or reject it? [Electronic version]. *Physics Education, 29*, 368-374.

AN ATTEMPT TO TEACH THE THEORY OF SPECIAL RELATIVITY TO STUDENTS OF UPPER SECONDARY EDUCATION

Kyriaki Dimitriadi[1]
University of Athens

Krystallia Halkia
University of Athens

Constantine Skordoulis
University of Athens

Abstract

This work presents a qualitative research designed to investigate whether the basic concepts of the Theory of Special Relativity (TSR) can be taught effectively to the upper secondary education students. For this purpose, a teaching sequence has been designed and relevant educational material has been developed. The research instrument was worksheets consisted of short stories written in a popularized code. Each story dealt with a particular concept of Special Relativity and was based on the difficulties students face according to bibliography. Thought experiments, paradoxes and experimental data were used to construct the stories. In the worksheets, students were invited to respond to the content of each story.The sample of the research was 20 students in the 10th grade. They worked in groups and their discourses were recorded. The results of this study show that in spite of the didactic difficulties of the theory (which is contrary to everyday experiences) students were able to grasp the basic ideas of TSR and to "construct" its axioms. They also dealt with its consequences (the relativity of simultaneity, time dilation and length contraction) up to a point but they tended to classify them as distortions of perception.

Introduction

Nowadays, the members of the science education community seem to agree that there is a need for updating the physics curricula in secondary education. A reason for this is that the physics concepts and theories covered in many physics curricula do not go beyond those at the beginning of the 20th century (Arriassecq et al, 2007). A major topic that has marked "modern physics" (physics of the 20th century) is the Theory of Special Relativity (TSR).

Rationale

One of the main reasons supporting the introduction of TSR in upper secondary education is its cultural value, since it helps students realize that physics is a human enterprise, develop their scientific awareness and enrich their science knowledge background with aspects of "modern" physics that have generated radical changes in science (Borghi et al, 1993; Arriassecq et al, 2007). In the relevant bibliography several alternative conceptions can be traced about the frames of reference, the speed of light and the relativity of simultaneity, time dilation and length contraction (Saltiel and Malgrange, 1980; Villani & Pacca, 1987; Panse et al, 1994; Scherr, 2001).

---

[1] Scholar of State Scholarships Foundation

The purpose of this study is to investigate whether and to what extent students of secondary education can understand the basic concepts of the TSR and what difficulties they face. Specifically, if they are able to "construct" the axioms of the theory and to deal with some of its consequences (the relativity of simultaneity, time dilation, length contraction). To this purpose, a didactic transformation of the scientific knowledge to school knowledge has taken place and educational material has been developed based on informal sources of science learning (popularized physics books, press science articles, DVDs about science).

Methods

Based on the findings concerning the alternative ideas of the analysis of the relevant literature and of a pilot study (Dimitriadi et al, 2005), a teaching sequence consisted of six sessions (meetings) was developed in order to introduce TSR to upper secondary education students. Each session lasted approximately 60 minutes. The research method used was that of *teaching experiment*, through which teaching and learning processes are investigated (Komorek & Duit, 2004). The sample was 20 students in the 10^{th} grade working in groups of 4.

Students were asked to work on worksheets consisting of simple short stories, containing everyday life experiences, thought experiments, paradoxes, experimental data, passages from popularized science books and other informal sources of science learning (newspapers, dvd). Each story deals with a particular concept of Special Relativity and it is based on the alternative ideas students face according to bibliography. Through these stories students were expected to reveal their alternative conceptions and to construct knowledge about the investigating concepts.

Members of each group had to discuss and argue about the viewpoints provided in these stories and answer the questions cited in the worksheets. The analysis of the collected data was based on the content analysis (worksheets) and on the discourse analysis (cassette recordings).

The results of this study suggest that it is possible to introduce TSR in upper secondary education. Students were able to follow the syllogisms, to discuss and to provide arguments.

As it was expected they faced some problems. Particularly, students met difficulties in grasping the relativity of motion and in using the frames of reference effectively. Most of them believed that motion is absolute and that there is a privileged frame of reference (the earth's or the sun's frame, depending on the problem). It was also obvious that they had not understood completely the Newton's laws of motion. For example, in one group they said that an object is moving only when it consumes energy, violating the law of inertia.

In addition, they tended to make a distinction between "real" motion, which has a dynamical cause, and "apparent" motion, which is an optical illusion as also Saltiel and Malgrange (1980) point out. For example, they said "Well, the train is really moving since it needs energy, whereas the platform seems that is moving to a passenger who is inside the train, but this doesn't really happen"

All students knew that the speed of light is finite, but they supported that surpassing this speed is a technological matter. The experimental evidence given to them afterwards, helped them overcome this difficulty. Though the invariance of the speed of light is beyond their perception, students were able to handle it.

The "Einstein's train" paradox helped them deal with the concept of the relativity of simultaneity, but two groups expressed the idea that actually only the stationary observer is right. This shows that it was difficult for them to apply the equivalence of the inertial observers. Sometimes, they talked about not only if the observer is moving or not but also if he is at the right point (which was in the middle of the distance according to them).

Table 1. The teaching sequence

Meeting	Topic	Conceptual steps	Educational Material used
1st	1st axiom of the TSR	• Motion is relative • Why we need the frames of reference • Practicing with different frames of reference. • Introduction of the concept of inertial frames of reference. • Equivalence of the inertial frames of reference.	Simple short stories concerning snapshots of everyday life
2nd	2nd axiom of the TSR	• Speed of light: finite • Speed of light: the maximum speed existing in nature • Invariance of the speed of light.	• Passage from a popular science book ("The time and space of Uncle Albert", Russell Stannard, 1989) • Dialogue among friends supporting alternative ideas. Experimental data. • Calculation of relative speeds using an example where the speed of light was surpassed and use of a passage from a popular science book ("Relativity visualized", Epstein, 1985)
3rd	Relativity of simultaneity	Thought experiment: Einstein's train paradox.	Visualization of the thought experiment (dvd: "Einstein, his life and his work")
4th	Time dilation	Calculation of time interval by two observers moving at a uniform speed to each other.	• Passage from a popular science book ("Space and time", Strnad & Podreka, 1993) • Experimental evidence that proves the time dilation (relativistic time dilation in muon decay) from a newspaper extract.
5th	Length contraction	Calculation of length separation by two observers moving at a uniform speed to each other.	An example from a popular science book ("Relativity visualised", Epstein, 1985)
6th	Evaluation	• Discussion about the differences between classical physics and TSR • Evaluation of the procedure.	Passage from a popular science book ("Mr Tompkins in Paperback", Gamow, 1990)

Results

Time dilation and length contraction were also difficult to grasp. They had no problem to follow the syllogisms and to come to the conclusions, but afterwards at the elaboration of the subject, they tended to classify these relativistic effects as distortions of perception. Interesting is the fact that it was much more difficult for them to deal with length contraction than time dilation. One student, for example, insisted that: "the length of a desk is specific. It doesn't matter how much we measure it. It is what it is now. We can see it." When he was asked, "Why didn't you say the same about the time?" he answered: "Well, time is another story. It is different. We cannot see time..." In this case, as Machold (1983) supports, the syllogisms needed are more complicated.

Students evaluated the whole procedure highly; though the course was not obligatory and they had to work after school, their attendance was regular and they were willing to discuss the subject after the end of the meetings.

Conclusions and Implications

"Special relativity offers an opportunity to channel student interest into a challenging intellectual experience" (Sherr, 2001). This study proves that in spite of the didactic difficulties of the theory (which is contrary to everyday experiences) students are able to grasp the basic concepts and ideas of the theory and to "construct" its axioms. In addition, they were able to deal with the consequences up to a point. Thus, further elaboration of the relevant material is needed to help students fully realize these consequences.

The use of informal sources of learning in the design of educational material was crucial, because they helped students comprehend easily the relevant material and deal with complicated science concepts. Moreover, the way the experimental implementation was applied has stimulated the students' interest and kept their motivation high.

References

Angotti J. A. P.,. Caldas I. L, Delizoicov Neto D., Rüdinger E., Pernambuco M. M. C. A. (1978). Teaching Relativity with a different philosophy, *American Journal of Physics* 46 (12).

Arriassecq, I. & Greca, I. (2007). Approaches to the Teaching of Special Relativity Theory in High School and University Textbooks of Argentina, *Science & Education 16: 65-86.*

Borghi, L., De Ambrossis, A. & Ghisolfi, E. (1993). Teaching Special Relativity in High School, *Third Misconceptions Seminar Proceedings.*

Dimitriadi, K., Halkia, K. & Skordoulis, C. (2005). Basic Concepts of Special Theory of Relativity in Secondary Education: Do students understand them?, e-proceedings of ESERA conference *"Contributions of Research to Enhancing Students Interest in Learning Science",* held in Barcelona, 28 August-1 September, 2005.

Epstein, L. (1985). *Relativity visualized,* copyright Lewis C. Epstein.

Erickson, F. (1998). Qualitative Research Methods for Science Education. *In International Handbook of Science Education.* Fraser B.J and Tobin K.G. (eds) pp. 1155-1173

Gamow, G. (1990). *Mr Tompkins in paperback,* Cambridge University Press

Komorek, M. and Duit, R. (2004). The teaching experiment as a powerful method to develop and evaluate teaching and learning sequences in the domain of non-linear systems, *International Journal of Science Education,* vol.26, no. 5, 619-633.

Machold, A. (1983). Zur qualitativen Behandlung der speziellen Relativitaetstheorie, *Der Physikunterricht 1,* s.39-49.

Panse, S., Ramadas J. and Kumar A. (1994). Alternative Conceptions in Galilean Relativity: frames or reference. *International Journal of Science Education,* 16, no1, 63-82.

Saltiel, E. & Malgrange, J. (1980). 'Spontaneous' ways of reasoning in elementary kinematics. *European Journal of Physics,* 1, 73

Scherr, R. (2001). *An Investigation of Student Understanding of Basic Concepts in Special Relativity.* Doctoral dissertation, University of Washington.

Stannard, R. (1989). *The time and Space of Uncle Albert.* Faber and Faber, London.

Strnad, J. & Podreka, E. (1993). *Prostor in cas, (In English: Space and time),* Mladinska Kljiga Zalozba, Slovenia.

Villani, A. and Pacca, L. (1987). Students' spontaneous ideas about the speed of light. *International Journal of Science Education,* 9, no1, 55-66.
Dvd: "Einstein, his life and his work" 1996, 2004 WGBH Educational Foundation. A NOVA production by Green Umbrella Ltd. for WGBH/Boston in association with BBC-TV and Sveriges Television.

THE RECONSTRUCTION OF THE FORMATION OF A MOUNTAIN RANGE BY STUDENTS AND THE PRINCIPLE OF ACTUALISM

Denise Orange Ravachol
CREN, EA 2661, Université de Nantes - IUFM des Pays de la Loire

Abstract

Our contribution studies the way students (13-14 and 16-17 years old) manage to reconstruct the formation of a mountain range, as one aspect of the reconstruction of the Earth's history, in order to see how these students can work spontaneously and what controls they use in their reconstruction of this phenomenon. To investigate these issues, our research is based on epistemological references (the importance of the methodological principle of actualism [uniformitarianism], the need for empirical traces) and on science educational works (students' relations to the geologic time). We analyzed students' written productions from two secondary school classes, collected at the beginning of the learning sequence. This study shows that students use a basic form of actualism and "storytelling" and highlights the conditions that allow them to move from a common-sense reasoning to a problematized approach to historical phenomena.

Introduction

Our contribution studies how students of secondary schools can reconstruct scientifically the history of the Earth. More precisely, how can they reconstruct the geological history of a mountain range? And how can they give up a common trend of explanation such as rudimentary stories?

Rationale

Many works in epistemology and science education can help us in this research.

- Geologists cannot build the history of the Earth in a strictly determinist framework or as a cosmogonic system (Hooykaas, 1963). This is not possible for three main reasons:

- Only modern and superficial geological processes or events are directly observable. Moreover, they are very complex.
- If geologists want to construct an "*actual*" past which is the opposite of purely imagined reconstitutions, they need empirical clues and traces of the Earth's past.
- The reconstruction of the past of the Earth must take into account the contingency that leads us to consider that other options might exist. As S.J. Gould (1989, p.278) wrote: "*Historical events do not, of course, violate any general principles of matter and motion, but their occurrence lies in a realm of contingent detail* ."

- In order not to just consider the idea of only a possible past, but to stay within the limits of a "*real*" past (an "*actual*" past), and in order to set certain sections of the past by operating or events needs, scientists are used to mobilizing the principle of actualism and empirical traces of the Earth's past.

- The methodological principle of actualism (uniformitarianism) represents a safeguard against too speculative and *ad hoc* reconstitutions. This assumption states that the present is the key to the past and, more specifically, that the phenomena at the origin of past geological changes (their geological causes) existed and still exist in the current nature (Gohau, 2003; Oldroyd, 2003). So this principle is "*the bridge that enables our imagination to go from the present to the past, to visualize with some confidence what no human eye has seen*" (Hooykaas, 1963, pp. XI-XII).

Geologists built thus a reasoned history of the Earth by articulating the present with the past and by mobilizing empirical traces of the past.

Several research projects already go deeper into the students' relationship with geologic time (particularly Dodick & Orion, 2003, Hidalgo & Otero, 2004). Based on the epistemological previous points of reference and on these works in science education, we want to characterize students' reasoning and their control over their explanation. How do low and high secondary school students at the medium level deal with the problems of historical geology? What controls do they exert over their reconstruction of Earth's phenomena or events? In this paper, we limit our study to the historical problem of the formation of a mountain range.

Methods

The formation of mountain ranges in the French curriculum

In the French curricula (table 1), the historical problem of the formation of mountain ranges (Alps, the Himalayas) is studied in the third level of the lower secondary general school (13-14 years old) and in the second and third levels of the upper secondary school (16-18 years old, scientific orientation).

Table 1. The content of the French curriculum

Geological knowledge	Students
"*The collision of continents generates deformations and leads to the formation of the mountains ranges.*" (programme, SVT, classe de quatrième, p. 9)	Grade 8 13-14 years old
"*The collision is the result of the convergency of two continental lithospheres. It usually follows a subduction and leads to the mountain range formation. Some aspects of the geology of the Franco-Italian Alps serve as example to illustrate these processes.*" (programme, SVT, classe de terminale S, p. 25)	Grade 12 17-18 years old

The mobilization of an elaborate form of the Principle of Actualism

The work on the problem of the formation of a mountain range deals with reconstructing a plausible scenario in the framework of plate tectonics. It involves an elaborate form of the principle of actualism (Orange Ravachol, 2005): as it is not possible to observe a mountain rising in a lifetime, we must consider a current process capable of generating reliefs and structures beyond human time and consequently a very long period from which geological mountain ranges originate (figure 1).

The work on this problem is also based on the mobilization of empirical traces whose research is complicated due to the fact that, as the evolution of the Earth cannot be conceived in a determinist framework, it is unrealistic to imagine many traces of the past, either because they have been destroyed or that they are difficult to identify (Cêlal Sengör, 2005). The construction of the real history of a mountain range takes the form of a set of construction with some processes which happen at different scales and with a subtle game of time.

What do students try to explain spontaneously? How do they use the principle of actualism? How do they deal with empirical traces?

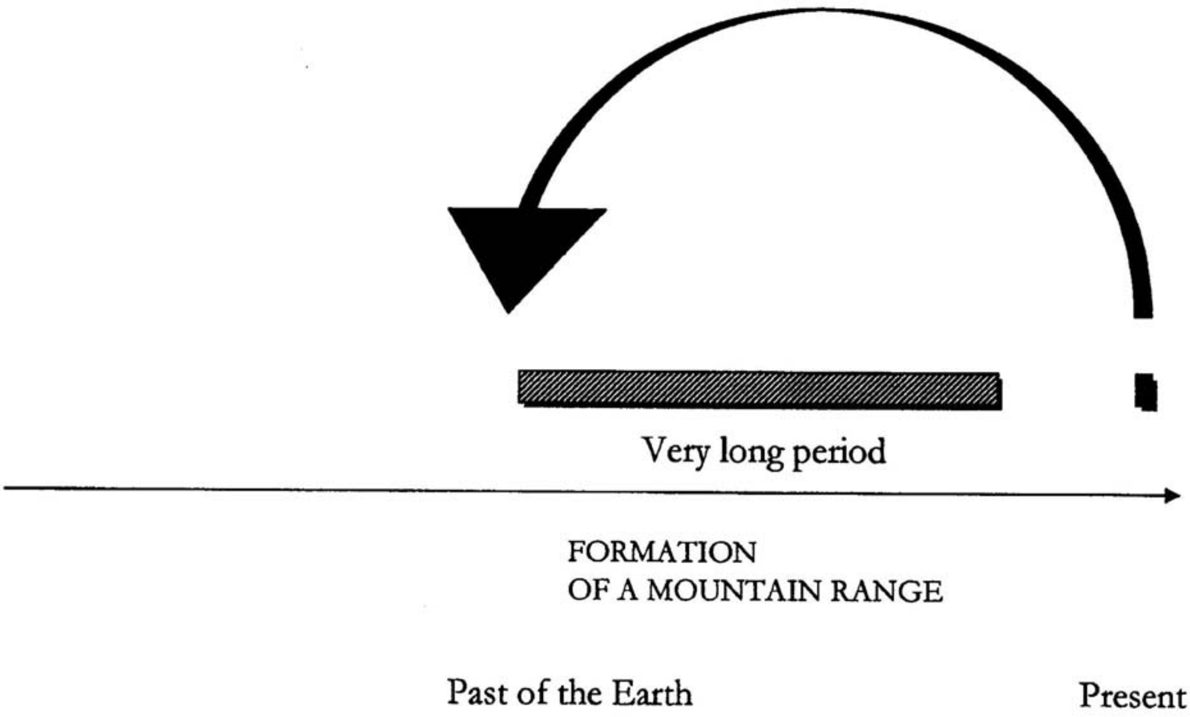


Figure 1. An elaborate form of the Principle of Actualism

A qualitative research

Our study is a qualitative study with a content analysis driven by the epistemological framework. Students' written productions (individual and groups; drawings and texts collected before the lesson) were analyzed by focusing on what they tried to explain (the shape of the landscape; organization of the rock formations; uplift) and their use of the principle of actualism (analogical actualism; elaborate form of actualism).

The experiment (table 2) was conducted at the beginning of the learning sequence and involved a class in the third level of the lower general secondary school (13-14 years old, 22 students) and a class in the second level of the upper general secondary school (16-17 years old, scientific orientation, 30 students).

- The 22 students (13-14 years old) had to explain how a mountain range (the Himalayas) was formed. They first worked individually and then were divided into 6 groups of 3-4 students.
- The 30 students (16-17 years old) searched individually an explanation for the formation of a part of Alps (*Le Chenaillet*).

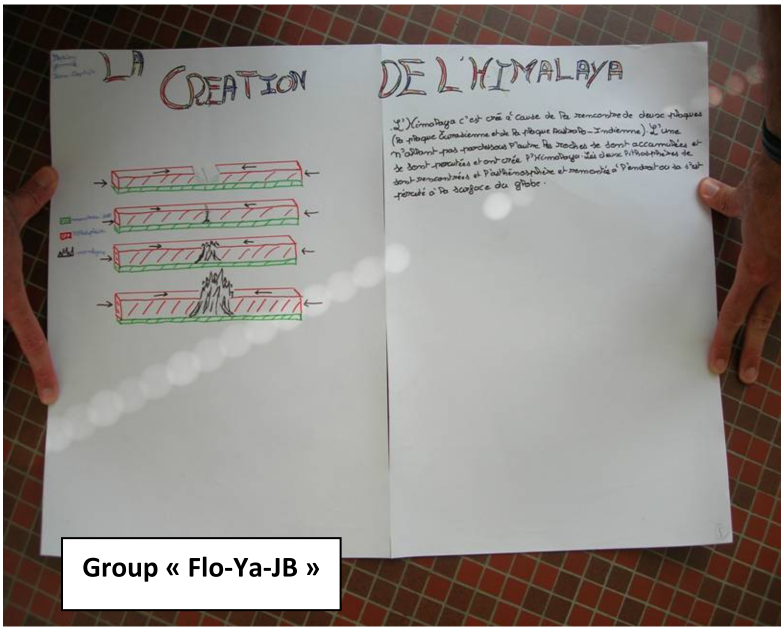
Table 2. The two studied classes

A class in the third level of the lower general secondary school (13-14 years old, 22 students) Explain how a mountain range (the Himalayas) was formed. Individual work, group	A class in the second level of the upper general secondary school (16-17 years old, scientific orientation, 30 students). Search for an explanation for the formation of a part of Alps *(picture of the Chenaillet).* Individual work

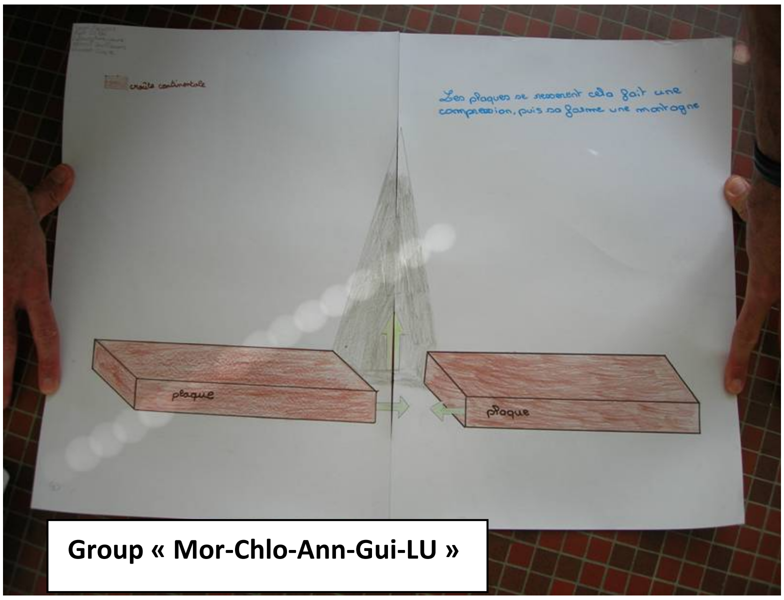
Results

Youngest students (13-14 years old)

The 22 young students spontaneously submitted an answer involving the meeting and confrontation of tectonic plates. They limited themselves to the explanation of the landscape, which acted as traces to be explained, by mobilizing a plate tectonics dynamic framework without restrictive rules, without needing many empirical traces, without time constraints. They were in the register of rudimentary stories ("storytelling"; figure 2).



« The Himalayas were created as the consequence of the encounter of two plates (the Eurasian and Australo-Indian plates). As one didn't go under the other, rocks accumulated and collided, and then created the Himalayas. The two lithospheres collided and the asthenosphere rose up on the place where the collision took place. »

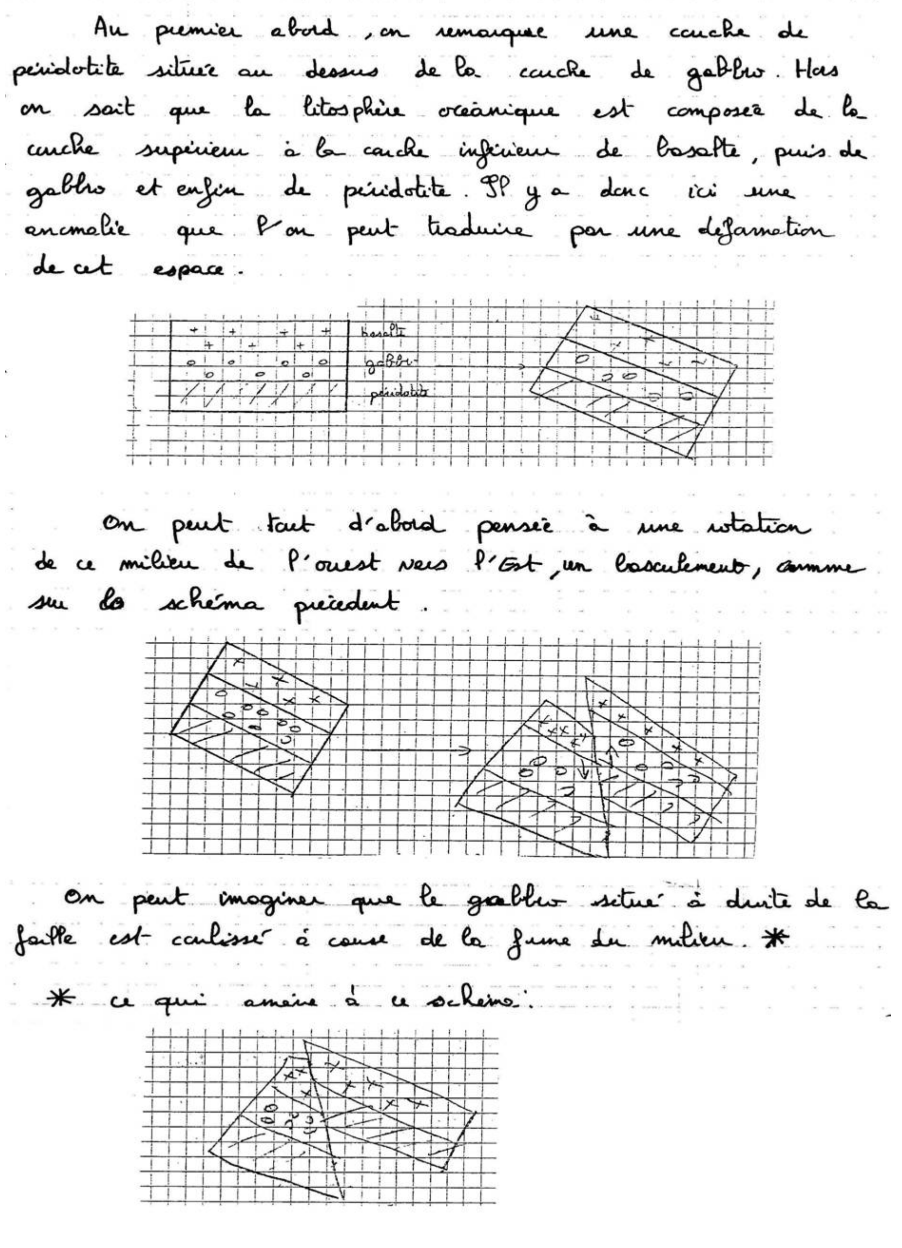


«The plates are tightening, so there is a compression, and then a mountain is generated »

Fig.2. Two "young students groups" ' written productions (13-14 years old)

Oldest students (16-17 years old)

With some knowledge of geological phenomena (for example some tectonic and orogenic processes), which was the case of the older students, we found that they submitted solutions quite close in appearance to those from scientist researchers. But they did not follow any rules or control to do so. Like handymen, they invented an assembly of the mountain range with some available objects (an ocean floor, plate tectonics, etc.) and some available tools (approximation, gauge, confrontation, reverse faults, etc.). The present period of nature was not really used as a reference, the need for long stretches of time from which phenomena originate was not built, and the historical contingency was far from being conceived. Students only offered solutions to obtain a geological outline, which was the evidence of an obvious lack of problematization. These students also resorted to "storytelling."

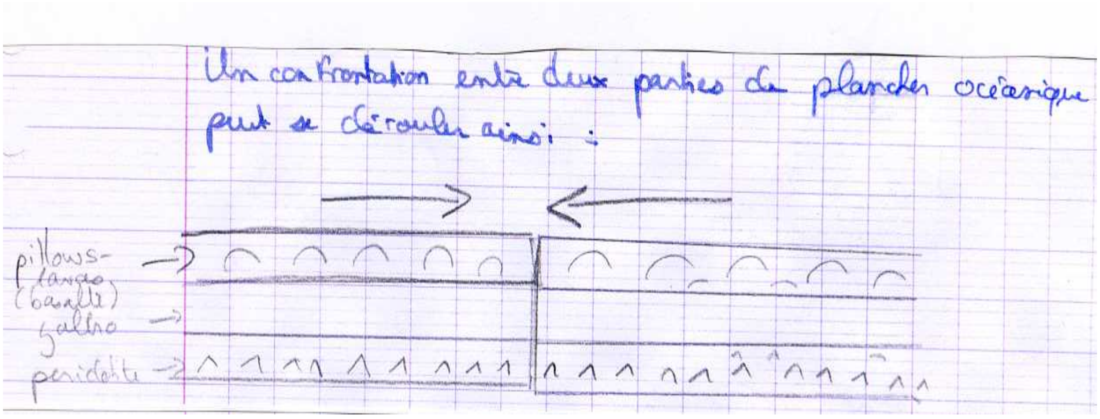


« We can notice a peridotite layer above the gabbro layer. But oceanic lithosphere is made with an upper layer of basalt, above a gabbro layer and then a peridotite layer below. Thus, there is an anomaly which can be interpretated as a deformation of this area. »

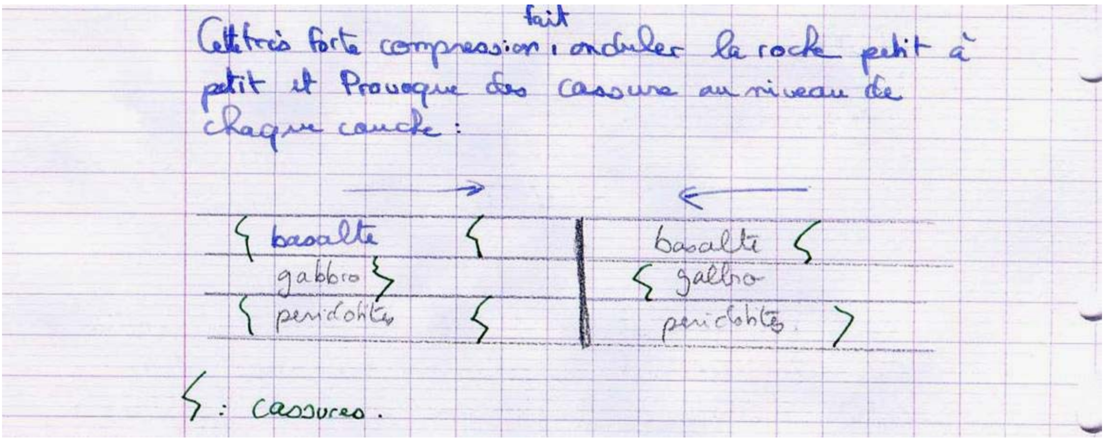
« We can first suppose that a rotation of this area from West to East, a flip, as shown on the previous figure.

« We can then suppose that the gabbro on the right of the fault had moved due to the shape of the area, giving the result shown by this figure.

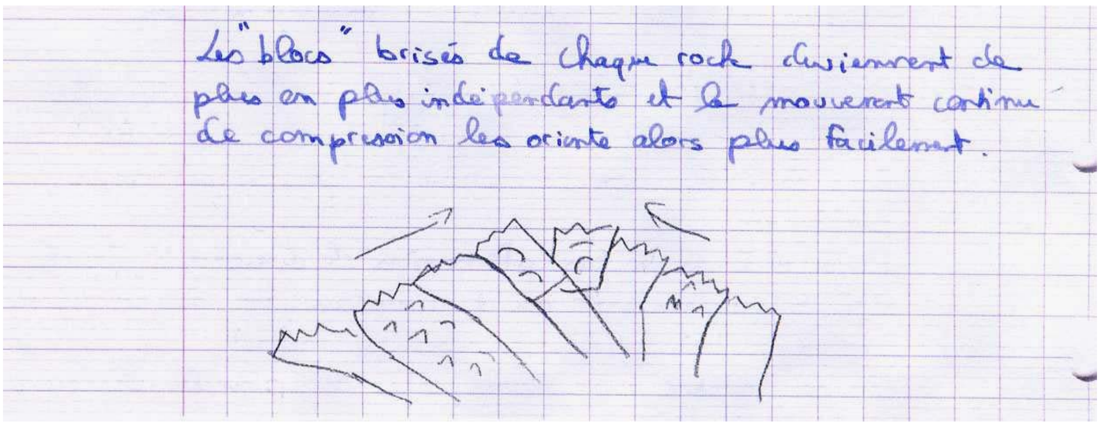
Figure 3. An oldest student's written production (Emmanuel, 16-17 years old)



« The confrontation between two oceanic plates can happen like this »



« This very strong crushing folds the rock and provokes crackles in every layer »



« Broken blocks of every rock become more and more independent. And the continuous crushing process can adjust them more easily.

Figure 4. Another oldest student's written production (Thomas, 16-17 years old)

Except for the relief, which is the empirical trace to be explained, students do not use spontaneously empirical traces. However, this use is better among oldest students (they have some knowledge of geological processes).

In short

The next table (table 3) sums up the formation of a mountain range as seen by young and oldest students.

Table 3. The formation of a mountain range as seen by students (13-14 and 16-17 years old)

Youngest students	Oldest students *(better geological knowledge)*
- Systematic use of the tectonic plates to explain the processes. - Limited use of empirical traces - No time and space rules	- Systematic use of the tectonic plates to explain the processes. - Better use of empirical traces - Still no time and space rules - « DIY geology » to explain geological features
« **Storytelling** »	

Conclusions and Implications

On historical geology problems, students do not work exactly as the scientists. They mobilize instinctively an analogical actualism and a catastrophism of convenience (without taking into account long periods of time). Dealing with the empirical traces, they are more focused on their recognition than on their building in a problematic situation. Does this mean that it is impossible for students to build the history of the Earth? We do not think so. Their resistance and their difficulties can help us to conceive devices to overcome them. So our research develops the three following propositions:

- Need for reinforcing the empirical and theoretical geological experience of students, which is useful for a reconstitution of the history of the Earth, so that they will be able to stand out for immediate solutions and will be able to consider some possible histories of the Earth (opening up of possibilities) but also will control these "possibles" by closing impossibilities.

- Allow students to control their geological stories, with a critical eye (debate allowing them to develop reasons, especially why some scenarios are impossible, and an elaborate form of the principle of actualism). As it happens among researchers, students would overcome difficulties to build an "actual" history of the Earth in those very far-off times by a step back from the critical notion of traces of geological processes and an involvement of an elaborate actualism.

- Allow students to understand that scientific knowledge is not limited to the solution of scientific problem. It is also important to know why this solution is possible and others impossible (the *« learning-through-problematization »* approach; Fabre & Orange, 1997).

These conditions might allow to remove the reconstruction of the Earth's history from the common sense and from an immediate empiricism leading to a rudimentary (hi)story and to consider it as a problematization from which reasons of a great History could emerge.

References

Cêlal Sengör, A.M. (2005). *Une autre histoire de la tectonique.* Paris : Collège de France/ Fayard.
Dodick, J. & Orion, N. (2003). Measuring student understanding of geological time. *Science Education*, vol. 87, Issue 5, pp.708-731.
Fabre, M. & Orange, C. (1997). Construction des problèmes et franchissements d'obstacles. *ASTER*, 16, pp. 37-57. Available on Internet: http://documents.irevues.inist.fr/handle/2042/8550)

Gohau, G. (2003). *Naissance de la géologie historique*. Paris : Vuibert-ADAPT.

Hidalgo, A.J. & Otero, J. (2004). An Analysis of the Understanding of Geological Time by Students at Secondary and Post-Secondary Level. Research Report. *International Journal of Science Education*, 4 June 2004, vol. 26, no.7, pp.845-857.

Hooykaas, R. (1963). The Principle of Uniformity in Geology, Biology and Theology. Leiden.

Gould, S.J. (1989). Wonderful Life, The Burgess Shale and the Nature of History. W. W. Norton & Company.

Lévi-Strauss, C. (1962). *La pensée sauvage*. Paris : Plon.

Oldroyd, D. (2003). The Earth Sciences. In *From Natural Philosophy to the Sciences* edited by David Cahan. University of Chicago Press, 2003.

Orange Ravachol, D. (2005). Problématisation fonctionnaliste et problématisation historique en Sciences de la Terre chez les chercheurs et chez les lycéens. *ASTER*, *40*, 177-204.

THE ROLE OF MICROCOMPUTER-BASED LABORATORY IN FACILITATING THE CONSTRUCTION OF KNOWLEDGE IN OPTICS

Fatma Caner
Marmara University, Turkey

Feral Ogan Bekiroglu
Marmara University, Turkey

Mehmet Aydeniz
University of Tennessee, Knoxville

Abstract

Traditional science teaching is not effective enough in altering student understandings of the physical science concepts. Research shows that students of all ages learn science better by participating actively in the critical thinking and by interpreting physical phenomena. Microcomputer Based Laboratory (MBL) is a technological tool composed of a computer, an interface device, probes, and sensors. MBL allows immediate real-time data collection and provides graphical representations of the collected data. Literature suggests that MBL data collection tools allows students to gather data more quickly so that they have enough time to ask and answer "what-if" questions. The purpose of this study was to explore and examine the role of MBL technology in students' conceptual understanding of optics. This study took place in an urban high school physics classroom in Turkey during five weeks. A multi-case study design guided this research with 9 participants. This study illustrated that the MBL technology is an effective instructional tool in stimulating students' learning of science. This technology promotes students' conceptual understanding of optics concepts as revealed through the difference between pre and post tests.

Introduction

Traditional science teaching is not effective enough in altering student understandings of the physical science concepts (Thornton, 1987). Literature reveals that students who learn science in traditional instructional settings often leave the classroom with misconceptions (Bernhard, 2003). Teaching science often takes place in the form of teacher telling, note taking and regurgitating information as response to stimuli provided by the teacher. McDermott (1992) states that teaching by telling is an ineffective method of instruction for most students and thus it is unsuccessful in helping students to overcome certain conceptual difficulties. Research shows that students of all ages learn science better by participating actively in the critical thinking and by interpreting physical phenomena (Donovan & Bransford, 2005; Gallagher, 2007). Students may be able to develop coherent conceptual understanding of science concepts through the use of well integrated instructional technologies such as micro-computer based laboratories (MBL). MBL is a learning technology that allows students to gather, analyze and interpret data (Bernhard, 2007). Research has revealed that the use of such technologies can encourage students to be active learners in the learning process (Sokoloff &Thornton, 2007).

Microcomputer Based Laboratory (MBL) is a technological tool composed of a computer, an interface device, probes, and sensors. MBL allows immediate real-time data collection and provides graphical representations of the collected data. Literature suggests that MBL data collection tools allows students to gather data more quickly so that they have enough time to ask and answer "what-if" questions (Kreuager & Rawls, 1998). According to Bernhard (2007), science instruments do not merely mirror reality but MBL gives students the opportunity to understand information on graphs and allows students how science can be done in the classroom. One of the benefits of MBL is that it allows for easy data collection, recording and graphing of data to support the construction of students' science concepts. Students spend less time in gathering data and more time in interpreting, discussing and evaluating the data when they engage in learning science through the use of MBL (Krajcik & Layman, 1992).

The use of instructional technologies influence students' attitudes towards learning of science .Thornton (1999) points out that students often perceive science as being difficult, boring and concerned with details. Students believe that science is exciting only for scientists because they are interested in discovery of something or development of models to explain the natural or physical phenomena (Osborne, 2004). Students develop such attitudes towards science primarily because they are ripped off from the opportunity to do science as their learning is often limited to listening to teacher lectures and note taking. Besides, students have a hard time connecting school science curriculum to their personal lives.

The integration of Microcomputer Based Laboratories is likely to influence students' attitudes towards science learning. That is because the integration of MBL into science instruction makes learning an active process and enables students a context to talk about science. For instance, students can make meaning of the data that they collect in real time (Sokoloff and Thornton, 2007).

Although Microcomputer- Based Laboratory (MBL) technology has widely been used in the U.S. and Europe, educators are just learning how to use such technologies in the classrooms in developing countries such as Turkey. The present study attempted to add to the literature by focusing the effects of MBL technology in the Turkish high school context.

Rationale

Science educators are increasingly becoming interested in determining the effects of MBL technology on students learning in science. Integration of technological tools such as MBL into science teaching and learning can serve multiple purposes. First, it can facilitate the enactment of reform objectives in science classrooms. For instance, it can provide the opportunity for students to understand the nature of science and scientific knowledge through inquiry made possible by the use of graphing calculators, computers, and data collection tools (Reid-Griffin, 2002). In addition to improved inquiry skills, the use of MBL in the classroom has been reported to contribute to increased student motivation and conceptual understanding (Brasell, 1987& Thornton, 1987).

Brasell (1987) presented that although students found kinematics to be a challenging subject, learning kinematics through MBL increased students' motivation. A study conducted by Friedler, Nachmias and Linn (1990) revealed that students succeeded in acquiring observation and prediction skills by learning science through MBL. Kelly and Crawford (1996) explored how MBL activities increased the quality of discourse and students' interaction during the discussions of motion concepts and graphics. Their findings suggest that the use of MBL in the classroom contributes to increased student participation in learning. Casey (2001) found that MBL increased the available time for analyzing, evaluating, and interpreting data; therefore, this technology gave more opportunities for problem solving, critical thinking, and reflection. Russel, Lucas and McRobbie (2002) stated that the use of MBL in the classroom motivated students for learning science and helped them to develop a better understanding of heat and temperature concepts. Nicolaou, Nicolaidou, Zacharia and Constantinou (2007) reported that MBL gave students opportunities to overcome a range of difficulties in interpreting graphs and improved their conceptual learning of melting and freezing.

The review of relevant literature cited above highlights the positive impact of the MBL technology on students' conceptual understanding of essential science concepts and their acquisition of scientific inquiry skills. The purpose of study was to explore and examine the role of MBL technology in students' conceptual understanding of optics.

Methods

A multi-case study design guided this research. Multi-case studies involve studying two or more subjects, settings, or depositories of data (Bogdan & Biklen, 1998).

Participants and Settings

This study took place in an urban high school physics classroom in Turkey. There were 23 ninth-grade students in the classroom and they had not had prior experience in learning science with MBL. The researchers taught the students how to use the MBL technology in an effort to implement the study with maximum fidelity. The students were asked to use the MBL equipment with various sensors during the orientation session. The study started only after the researchers were convinced that all the students knew how to use the MBL technology. Nine students were randomly asked to become participants in the study. All nine of them agreed to become participants. They included seven males and two females. Their grades in physics ranged from 1 to 5, with 5 being the highest grade. Table 1 presents some demographic data about the participants.

Table 1.Demographics of participants

Participants	Age	Gender	Interest of computer	Grade of exam out of 5
P-1	14	F	Not interested in computer	1
P-2	14	M	Highly interested in computer	5
P-3	14	F	Interested in computer	5
P-4	13	M	Highly interested in computer	4
P-5	14	M	Interested in computer	4
P-6	14	M	Not interested in computer	2
P-7	14	M	Interested in computer	3
P-8	13	M	Highly interested in computer	3
P-9	14	M	Not interested in computer	4

Role of the Researcher

The first author had worked with the teacher of the classroom as a mentor before this project started. The researcher had assisted to the teacher with planning and teaching of the laboratory sessions for two semesters prior to the study. The first author had built a good rapport with the students because of her internship. This relationship helped with the recruitment of the participants.

Instructional Context

This research lasted five weeks for the period of teaching Optics Unit. The teacher of the classroom was the first author throughout the research; thus, she planned the whole instruction, guided the students and prepared worksheets by following the principles of open-inquiry model of instruction. The students worked in groups during the lab activities. They used light sensors as a part of the MBL technology and followed the predict-observe-explain procedure to complete the learning activities. The activities lasted approximately 90 minutes and were related to the concepts of light, brightness, color, illumination, and polarization.

Data Collection

Data were collected by using various sources including semi-structured interviews, the researcher's field notes, the participants' written documents (i.e. worksheets, assignment sheets, pre- and post-test data), and classroom discourse. The students' conceptual understanding of optics was measured through a set of open-ended questions through the pre- and post-tests. The pre- and post tests were identical. There were 10 open ended questions in the tests. The purpose of the semi-structured interviews was two-fold: knowing the participants better by asking questions about their daily lives, families, and hobbies; and determining the role of MBL technology in their learning. The semi-structured interviews were video-taped. Other data sources were used to detect student learning as well as to understand the role of MBL in their learning. Three video cameras were placed in the room to record the classroom environment. The students' conversations in their groups were tape-recorded. In an effort to achieve triangulation, multiple data sources were used and the data were analyzed by two coders (Maxwell, 1996). Triangulation was used to develop valid and meaningful propositions and to minimize the bias inherent in qualitative research (Mathison, 1988).

Data Analysis

Data analysis took place in multiple steps. First, the learning objectives assessed by the questions in the pre- and post-tests were identified. Second, the concept of illumination was represented with 1; therefore, all the indications for the objectives that were related to illumination concept began with 1. Similarly, the concept of light intensity was represented with 2, the concept of color was represented with 3, and the concept of polarization was represented with 4. Table 2 shows how the Optics concepts are covered by the learning objectives. For instance, calculation of illumination intensity (indicated as 1A), interpreting how illumination changes according to angle (indicated as 1B), and understanding mathematical model of changes in illumination intensity according to angle (indicated as 1C) were asked under three different questions. In order to indicate that these objectives were assessed by three different questions, they were coded as 1Aa, 1Ab, 1Ac; 1Ba, 1Bb, 1Bc; and 1Ca, 1Cb, 1Cc. In addition, comprehending working principle of a photometer (indicated as 2A) and understanding of polarization (indicated as 4A) were asked under two different questions. Therefore, these objectives were coded as 2Aa, 2Ab and 4Aa, 4Ab. Besides, understanding the effect of mirror in increasing illumination (indicated as 1D), interpreting how illumination intensity changes according to distance (indicated as 1E), understanding mathematical model of changes in illumination intensity according to distance (indicated as 1F), calculation of light intensity (indicated as 2B), understanding how light colors occur (indicated as 3A), and realizing properties of Polaroid sunglasses (indicated as 4B) were asked under one questions. Hence, the codes for these objectives were remained same.

After the learning objectives were established, the results of pre- and post-tests were compared to determine improvements in the students' conceptual understanding of optics. Qualitative analysis involved verbatim transcripts of the tape and video recordings and written documents. The data gathered from worksheets, assignment sheets and classroom discourse were analyzed inductively to identify the learning objectives that the participants achieved during their involvement with MBL. If the participant achieved a learning objective in all of the related questions of the post-test that s/he did not achieve in the pre-test and this objective was identified in either discourse or worksheets or assignment sheets that belonged to the participant, it meant that s/he had learned the content represented in the objective with the help of MBL technology. The interview data were used to support the researchers' interpretations about the role of MBL in students learning by receiving the participants' opinions.

Results

Table 3 demonstrates all participants' results but we will only elaborate on one case in our discussion due to page limitations. This table was illustrated by using data, which was gathered from pre- and post-tests, assignments, class activities and discourse in the class. Also these data such as assignments, activities and discourse were the evidence in order to show that learning objectives which were achieved in post-test were learnt by MBL or not.

Table 4 shows the performances of only one participant (P-1) in the pre- and post-tests as well as her performances in the assignments, activities, and classroom discourse. The participant's performance was matched with the learning objectives covered during this study. The learning objectives indicated in bold illustrate that the MBL technology had a role in the student's understanding of the objectives. As indicated in Table 4, the participant knew only three learning objectives that is called interpreting how illumination changes according to angle (1Bb), understanding mathematical model of changes in illumination intensity according to angle (1Cb) and comprehending working principle of a photometer (2Ab) before the instruction as revealed through her pre-test results. After the instruction, she knew eight learning objectives and five of them were the new ones.

Table 2. Learning Objectives and their indications

Concepts	Codes based on the Frequency	Learning Objectives	Indications
Illumination	1Aa, 1Ab,1Ac	Calculating illumination intensity	1A
	1Ba,1Bb, 1Bc	Interpreting how illumination changes according to angle	1B
	1Ca,1Cb,1Cc	Understanding mathematical model of changes in illumination intensity according to angle	1C
	1D	Understanding the effect of mirror in increasing illumination	1D
	1E	Interpreting how illumination intensity can change according to distance	1E
	1F	Understanding mathematical model of changes in illumination intensity according to distance	1F
Light Intensity	2Aa, 2Ab	Comprehending working principle of a photometer	2A
	2B	Calculate light intensity	2B
Color	3A	Understanding how light colors occur	3A
Polarization	4Aa, 4Ab	Understanding of polarization	4A
	4B	Realizing properties of Polaroid sunglasses	4B

Table 3. All participants' results

Participant	Learning objectives in pre-test	Learning objectives in post-test	Learning objectives in assignments	Learning objectives in activities	Learning objectives in discourse
P-1	1Bb, 1Cb, 2Ab	1Aa,1Ab, 1Ac, 1Ba, 1Bb, 1Bc, 1Ca, 1Cb, 1Cc, 1D, 1F, 2Ab, 3A, 4Aa,	3A	1A,1E, 3A,4A	1 E, 1C, 1A, 2A,4A
P-2	1Ab, 1D 2B,	1Aa,1Ab, 1Ac, 1Ba, 1Bb, 1Bc, 1Ca, 1Cb, 1Cc, 1D, 1E, 1F, 2Aa, 2Ab, 3A, 4Aa,4B,	3A	1A,1E, 3A,4A	2A,3A,4A
P-3	1Ab,1Ac,1Bc,1Cc 1D	1Aa,1Ab, 1Ac, 1Ba,1Bb, 1Bc, 1Ca, 1Cb, 1Cc,1D,1E,2Ab, 3A,4Aa		1A,1E,3A, 4A,	1B, 1D,2A,
P-4	1Aa,1Ba, 1Ca 3A	1Aa,1Ab, 1Ac, 1Ba, 1Bb, 1Bc, 1Ca, 1Cb, 1Cc, 1D, 1E, ,2Aa, 2Ab 3A, 4Aa,4Ab	3A 4A 4B	1A,1E, 3A,4A,	4A,1A,1B, 3A,2A
P-5	1Ab,1Ac, 1D	1Aa,1Ab, 1Ac, 1Ba, 1Bc, 1Ca, 1Cc, 1D, 1E, 3A, 4Aa,4Ab	3A	1A,1E, 3A,4A,	1E 3A
P-6	1Aa, 1Ca, 1D, 3A	1Aa,1Ab, 1Ac, 1Ba, 1Bb, 1Bc, 1Ca, 1Cb, 1Cc, 1D, 2Aa, 2Ab, 3A,4Aa,4Ab		1A, 3A,4A	2A,
P-7	1Aa,1Ab, 1Ac, 1Ba, 1Bb, 1Bc, 1Ca, 1Cb, 1Cc, 1D, 1E, 1F,2Aa, 2Ab, 3A,	1Aa,1Ab, 1Ac, 1Ba, 1Bb, 1Bc, 1Ca, 1Cb, 1Cc, 1D, 1E, 2Aa, 2Ab, 3A, 4Aa,4Ab,		1A,1E, 3A,4A,	4A
P-8	1Ab, 1D	1Aa,1Ab, 1Ac, 1Ba, 1Bb, 1Ca, 1Cb, 1D, 1E,2Ab,, 3A, 4Aa, 4Ab,		1A,1E, 3A,4A,	2A, 3A,4A
P-9	-	1Aa,1Ab, 1Ac, 1Ba, 1Bc, 1Ca, 1Cc, 1D, 1E, 3A, 4Aa	3A	1A,1E, 3A,4A,	4A, 1E

Table 4. Learning objectives achieved by P-1

Participant	Learning objectives achieved in pre-test	Learning objectives achieved in post-test	Learning objectives achieved in assignments	Learning objectives achieved in activities	Learning objectives achieved in discourse
P-1	1Bb, 1Cb, 2Ab,	**1Aa**, **1Ab**, **1Ac**, 1Ba, 1Bb, 1Bc, **1Ca**, 1Cb, **1Cc**, 1D, 1F, 2Ab, **3A**, **4Aa**	3A	1A, 1E, 3A, 4A	1A, 1C, 1E, 2A, 4A

The meanings of a, b, and c are that the same objective that was assessed in different questions.

P-1's physics grade was 1 out of 5. However, she was very active during the study than she usually had been in the class. P-1 improved her knowledge of optics by participating in the MBL activities. According to Table 4, MBL technology helped her to learn how to calculate illumination intensity (1A), the mathematical model of changes in illumination intensity according to angle (1C), how light colors occur (3A), and the concept of polarization (4A). However, it is interesting to see that although P-1 could interpret how illumination intensity can change according to distance (1E) during the classroom discourse and in her activity sheet, she could not reach this objective in the post-test. The inconsistency between her performance at the post-test and her classroom participation indicates that she might not reflect her learning during her post-test performance because of the exam anxiety or some other reasons that we could not detect.

She pointed out why her performance in the activities and her grade in Physics were not coherent with each other in the interview.

> Researcher: Why did you get "1" from the exam? Do you not like Physics?
>
> P-1: No. I do not know why.
>
> Researcher: Your performance was good during the class activities. You took part of the conversations and discussions. I was pleased of your performance.
>
> P-1: Yes, the activities were fun. I liked them very much. During the activities, I had a chance to perform myself. I enjoyed and I participated in most of the activities. I liked the discussions. I am surprised by my performance in the class.

MBL gave P-1 the opportunity to perform well by taking part of the class activities and by discussing the results of the experiments with her partners. Furthermore, she was eager to do science in the class. Using MBL in the science classroom helps to give students the autonomy to develop and realize their own ideas. Students feel that they are active participants in the learning process. This feeling of success may make them interested in Physics or doing of science.

She also expressed the positive effects of MBL technology by saying that:

> MBL technology helped me to understand many things. Visuality, real-time data collection and graphics made it easy for me to learn what was covered in the class…Because we discussed the results, I realized what my mistakes were (P-1). Briefly;

- Before the instruction, P-1 could not achieve any of the 1A objectives. After the instruction, P-1 achieved all of 1A objectives.
- Before the instruction, P-1 achieved only one of three questions of 1C objectives. After the instruction, P-1 achieved all of 1C objectives.
- Before the instruction, P-1 achieved one of two questions of 2A objectives. After the instruction, P-1 could not achieve all of them because of some mathematical mistakes.
- Before the instruction, P-1 could not achieve 3A objective. After the instruction, P-1 achieved 3A objective.
- Before the instruction, P-1 could not achieve any of the 4A objectives. After the instruction, P-1 could not achieve all of them (one of two questions). It is realized that P-1 made a mathematical mistake and she could not answer the one of two questions of polarization correctly.
- Although P-1 achieved 1E learning objectives in the classroom discourse and activities, P-1 did not achieve this objective in the post-test.

The results from the between-case analyses pointed out that almost all participants learned how to calculate illumination intensity (1A), how light colors occur (3A), and the concept of polarization (4A) by using MBL. This we believe was a result of students' use of MBL. Moreover, MBL provided students with understanding of how illumination intensity changes according to angle (1B), how illumination intensity changes according to distance (1E), the working principle of a photometer (2A), properties of Polaroid sunglasses (4B), and how to calculate light intensity (2B).

From the participants' point of view, the benefits of MBL were as follows:

- opportunity to test hypotheses,
- ability to present real-time graphs on the computer's screen,
- opportunity to finish experiments easily in a short time,
- to allow plenty of time for interpretation and discussion,
- opportunity to repeat the experiment many times, and
- opportunity to share ideas.

According to the participants, MBL was fun. They also felt like scientists while completing the lab activities. This can be seen as indication of students developing positive attitudes towards science. The participants revealed that they would never forget the knowledge they acquired through participation in MBL learning activities.

Conclusions and Implications

Several conclusions can be drawn from the results of this study. First, the results show that the MBL technology is an effective instructional tool in stimulating students' learning of science. This technology promotes students' conceptual understanding of optics concepts as revealed through the difference between pre and post tests. Second, the MBL technology can play an important role in promoting students' interest in science and in physics particularly. Third, MBL promotes students' acquisition of scientific inquiry skills. For instance, it provides the chance for students to interpret real-time graphs provided by MBL. It offers the opportunity for students to manipulate variables and understand the relationships between them. MBL tools support quick data collection and data are presented graphically in real time, which in turn gives the students the opportunity to spend time on interpreting the results. However, in order for MBL to bring about real learning, student learning must be scaffolded by the teacher. These results are in line with the findings of Brasell (1987), Thornton and Sokoloff (1990), Reid-Griffin (2002), Russell, Lucas, and Mc Robbie (2004). This multi-case study adds to the literature investigating the role of Microcomputer-Based Laboratory in development of students' science content knowledge. Future studies should look at how the use of MBL may affect the learning of students with different achievement levels. Similarly, future studies can focus on teachers' use of MBL technologies in the classroom.

References

Bernhard, J. (2003). Pyhsics learning and microcomputer based laboratory (MBL): Learning effects of using MBL as a technological ans as a cognitive tool. In D. Psillos& K. P. &V. Tselfes & E. Hatzikraniotis &G. Fassoulopoulos & M. Kallery (Eds.), *Science Education Research in the Knowledge Based Society*, 313-321.

Bernhard, J. (2007). Thinking and learning through technology-Mediating tools and insights from philosophy of technology applied to science and engineering education. *The Pantaneto Forum*,27.

Bogdan, R. C., & Biklen, S. K. (1998). *Qualitative research in education: An introduction to theory and methods.* Needham Heights, MA: Allyn & Bacon.

Brasell, H. (1987). The effect of real-time laboratory graphing on learning graphic representations of distance and velocity. *Journal of Research in Science Teaching,* 24(4) 385-395.

Casey, D. A. (2001). *A cultural study of a science classroom and graphing calculator-based technology.* Unpublished Doctoral Dissertation, Virginia Polytechnic Institute and State University, Blacksburg, VA.

Donovan, M. J., & Bransford, J. D. (2005). *How students learn: Science in the classroom.* Washington, DC: National Academy Press.

Friedler, Y., Nachmias, R., & Linn, M.C. (1990). Learning scientific reasoning skills in Microcomputer-Based Laboratories. *Journal of Research in Science Teaching*, 27(2), 173-191.

Kelly, G.J. & Crawford, T. (1996). Students' interaction with computer representations: Analysis of discourse in laboratory groups. *Journal of Research in Science Teaching*, 33, 693–707.

Kreuger, A. &Rawls, G. (1998). Connecting points: teacher decision-making about student data-collection technology. *Journal of Science Education and Technology,* 7(3).

Layman, J. &Krajcick J.(1992). The microcomputer and practical work in science Laboratories',in Layton, D.(ed), *Innovations in Science and Technology Educatio,*Vol IV,UNESCO.

Mathison, S. (1988). Why triangulate? *Educational Researcher*, 17(2), 13-17.

Maxwell, J. A. (1996). *Qualitative research design: An interactive approach.* Thousand Oaks,CA: Sage Publications.

Mc Dermott L. (1992). How we teach and how students learn- a mismatch?. Proceedings ICPE 4th International Conference, Badajoz, Spain.

Nicolaou, C., Nicolaidou, I., Zacharia Z., & Constantinou C. (2007) Enhancing fourth graders' ability to interpret graphical representations through the use of Microcomputer-Based Labs implemented within an inquiry-based activity sequence. *Journal of Computers in Mathematics and Science Teaching*, 26(1), 75-99.

Reid-Griffin, A. R. (2002). *Observing middle school students' use of technology as a tool for learning.* Unpublished Doctoral Dissertation, North Carolina State University, Raleigh, NC.

Russell, D., Lucas, K. B., & McRobbie, C. J. (2004). Role of microcomputer-based laboratory display in supporting the construction of new understandings in thermal physics. *Journal of Research in Science Teaching*, 41(2), 165-185.

Sokoloff, D. R. & Thornton, R. K. (2007). Active learning with interactive lecture demonstrations. Proceedings of the ICPE, In Press.

Thornton, R. K. (1987). Tools for scientific thinking: Microcomputer – based laboratories for physics teaching. *Physics Education*, 22(4), 230-238.

Thornton, R. K., & Skoloff, D. R. (1990). Learning motion concepts using real-time microcomputer-based laboratory tools. *American Journal of Physics,* 58(9), 858-867.

Thornton, R.K. (1999). Why Don't Physics Students Understand Physics? Building a Consensus, Fostering Change. Chapter in *The Thirteenth Labor, Improving Science Education*, E.J. Chaisson, TC Kim ed., (Amsterdam, Gordon and Breach Publishers).

ACTIVITY THEORY AND LEARNING IN SCIENCE EDUCATION LABORATORY LESSONS. THE CASE OF MAGNETISM

Xarikleia Theodoraki
University of Ioannina, Greece

Katerina Plakitsi
University of Ioannina, Greece

Abstract

This is a study on connecting Science Education with Cultural Historical Activity Theory (CHAT). It is a case study, which focuses on a sequence of science education lab activities on the topic of Magnetism, organized by the University of Ioannina in Greece. When discussing activity, activity theorists are not simply concerned with "doing" as a disembodied action but are referring to "doing in order to transform something," with the focus on the contextualized activity of the system as a whole. We focus on object which even is constantly in transition and under construction, and "it manifests itself in different forms for different participants and at different moments of the activity". The Laboratory Lesson of Magnetism based on Activity Theory (LLMAT) is an experimental course in the University of Ioannina. We divided our research in seven steps: Organization of the LLMAT, Video studies, Evaluation, Questionnaires and Interviews, Elaboration of data, Students' Practice, and Reflection. In discussion we interpret objects' function. Other connected factors are subjects, tools, rules, community and division of labor. This study, even limited, could push forward the boundaries of teachers' training in science education. Especially it highlights teachers' capability to provoke meaningful learning in school science classrooms.

Introduction

This study belongs to a research program in the field of connecting Science Education with Cultural Historical Activity Theory (CHAT). It is a case study which focuses on a series of science education lab activities on the topic of Magnetism, organized by the University of Ioannina in Greece.

Activity theory is a psychological and multidisciplinary theory with a naturalistic emphasis that offers a framework for describing activity and provides a set of perspectives on practice that interlink individual and social levels (Engeström, 1987, Nardi, 1996, Roth & Lee, 2004). Activity theory is a theoretical framework for analyzing human practices as developmental processes with both individual and social levels interlinked at the same time (Kutti, 1996, Nardi, 1996, Kaptelinin & Nardi, 2006). This framework uses 'activity' as the basic unit for studying human practices.

When discussing activity, activity theorists are not simply concerned with "doing" as a disembodied action but are referring to "doing in order to transform something," with the focus on the contextualized activity of the system as a whole (Engeström, 1987, Kuutti, 1996). These relations between participant and object are not direct; rather, they are mediated by various factors, including tools, community, rules, and division of labor. From an analytical point of view it is quite complicated to handle this. To overcome this obstacle we adopted Engestrom's triangular model (1987). We focus on object which even is constantly in transition and under construction, and "it manifests itself in different forms for different participants and at different moments of the activity" (Hasu & Engeström, 2000).

Our research, concentrates on transferring the activity theory into some fields of science education. As it is combined with other relevant case studies it finally aims to validate the activity theory as an evaluation tool of scientific activities in different learning environments, as it is the school classroom, the laboratory etc.

Rationale

Activity theory is a theory with expanding applications in different fields of science. The students' group, the intermediary tools and the objects play an important role for establishing new principles or ideas in the context of rules that the entire group follows. The unit of analysis is the activity. This makes moving from one activity to another flexible, getting advantage of previous knowledge. Thus, the construction of knowledge becomes meaningful for the students who interact with one another as well as with tools and means into the community of learners and also within the activity's context (Engeström, 1999).

In this paper, we conduct a pilot research on testing activity theory as an analytical tool for science education activities. At first, we present a succession of laboratory activities concerning the magnetic properties of certain materials. In the second place, we analyze the framework of activity theory. The discussion concerns some limits of the transfer of activity theory to science education context (Cole, 1995). The research is based on a case study, using multiple methods for gathering data (Yin, 1994). We construct the LLMAT lab lesson (Laboratory Lesson of Magnetism based on Activity Theory), which is an experimental course in the Department of Early-Childhood Education in the field of science education. We use this lab lesson as a methodological tool to prepare university students to teach the magnetism issue in the classroom. The central point of LLMAT was our position about the educational success or failure which had been explicated as a collective activity in the social framework. The organization of LLMAT included the following steps:

Awareness. In this stage we co-decided that the topic of magnetism was the most interest topic for our studies.

Comparison. We compared the laboratory lesson of magnetism with the previous laboratory lessons and we used obtained knowledge, skills and materials.

Exploration/Activating Prior Learning. In this stage we tried to trace some students' obstacles on their conceptualization of magnetism.

Creation. We developed additional insights about the magnetism and we created activities based on our students' needs.

Students use their prior knowledge on this issue so as to categorize some materials by testing their behavior when they come close to a magnet. Then, they obtain further knowledge about magnetic properties and create new activities based on their prior knowledge. Furthermore, they take into account pupils cognitive obstacles in organizing the new activities so as to achieve the aim.

Methods

The study is based on Engeström's theory of expanding learning. The ideas presented in AT enhance and extend the practical concerns of tool usage, which are traditionally addressed by linking the design solution to socio-cultural and psychological aspects of the tool user. This approach highlights the importance of the tool user's cultural behavior revealed during tool usage. It seems that by analyzing human activity in context, using this framework, the computer tool developer can fully account for the complex and intertwining issues that impact on the usefulness of the computer tool through its design. Although the ideas presented in this framework sound promising by providing a much-needed common vocabulary for describing human activity, there is no standard method for putting Activity Theory ideas into practice (Nardi, 1996). The lack of a standard method for applying

Activity Theory could be attributed to the fact that there are several basic principles of Activity Theory (Kaptelinin, 1996) on which one could base their analysis. In addition, the framework itself is continuously evolving, in the sense that concepts from this framework have been interpreted and applied in various ways in different contexts. As a result, difficulties have come up concerning replicating, comparing and criticizing the approaches taken to apply Activity Theory.

The activity triangle model incorporates the Subjects, Object, and Community components; also mediators of human activity, namely: Tools, Rules and the Division of Labor. These components are discussed below.

The 'Object' component reflects the motivational or purposeful nature of human activity that allows humans to control their own behavior. Human activity is targeted towards the satisfaction of identified objectives. For this reason,, in this paper, the term 'object' is used in the "objective" (see Leont'ev, 1981, Pages 46-69) sense, so as to emphasize the purposeful nature of human activity. The 'Subjects' component of the model portrays both the individual and social nature of human activity as reflected through collaborations and consultations in order to satisfy a shared objective. The subjects' relationship with the object or objective of activity is mediated through the use of tools. The 'Tools' component of the model reflects the meditational aspects of human activity through the use of both physical and conceptual tools. Physical tools are used to handle or manipulate objects whilst conceptual tools are used to influence behavior in one way or another. The 'Community' component of the model puts the analysis of the activity being investigated into the social and cultural context of the environment in which the subject operates. This notion reaffirms the suitability of AT to the study of human practices in an organization. The Rules component highlights the fact that within a community of actors, there are bound to be rules and regulations that affect in one way or another the means by which activity is carried out. These rules may either be explicit, or implicit, for example, cultural norms that are in place within a particular community. The Division of Labor component refers to the allocation of responsibilities and variations in job roles of the subjects as they carry out activity in the community.

From a sample of 80 third-year students, three pairs were assigned to teach magnetic properties of various materials in three different pre- primary classrooms during their two-week practice in schools. In order to do this, they had to follow the steps of the LLMAT lab lesson practiced at University. At this point, we have to mention that students had little previous experience in a real classroom situation and they had to work with 18-20 pupils in class. Finally, teachers of the schools were willing to collaborate with the students during the whole procedure. Students participated in the designing of lab activities, following the eight- step model of Mwanza 2001(Mwanza, 2001).

1. Activity of interest. In this stage students modified the sort of activity in which they were interested in. The study of magnets began with an exploration of magnetic and non-magnetic properties of certain materials. We provided each pair of students with a horseshoe magnet and encouraged them to go on a "magnetic hunt" with their partner. They explored the room, predicting the behavior of different materials when they came close to a magnet.

2. Objective of activity. Students provided the reason the activity takes place. They shared their findings and made observations while experimenting with the magnets

3. Subject in this activity. They discussed about who were involved in activity (students, teachers, parents).

4. Tools mediating activity. Books and other materials were the tools by which the subjects (students) carried out the activity. At this point,, students demonstrated a deeper understanding of magnets and magnetic properties, knowledge of the earth's magnetic field and the way a compass works, and usage of magnets in everyday life.

5. Rules and regulations mediating the activity. They collectively accepted the rules that all had to follow during the activity: a) Each pair of students worked together to explore the strength of the magnet. b) They recorded their

findings in their datasheet. c) When students completed their experimentation, they discussed their findings in the classroom. d) Teacher evaluated the students' responses.

6. Division of labor mediating the activity. **Teacher**: 1) demonstrated how to set up the experiment (without actually demonstrating the results; this is for the students to discover), 2) brought the class together for discussion (What happened? Were our hypotheses correct? What conclusions can we make? Which poles repel or attract). **Students**: 1) worked in pairs. 2) experimented with a partner to determine the magnetic and the non magnetic properties of materials.

7. Community in which activity is conducted. In this step we defined the environment in which the activity was carried out. More specific, the environment was the class with Greek pupils.

8. Outcomes. This is the final step in which we provided an estimation of the results from carrying out this activity.

Results

During our collaboration with students we stressed the importance on developing, implementing, and studying science teaching and learning that is guided by several principles. We strived to offer both pupils and university students, multi-modal opportunities to engage with different levels of scientific activity, theorizing about the world around us and collecting and processing data (empirical evidence) either in the form of observations or designing experiments (Roth, Tobin 2007). In this section, we describe a hands-on task in one pre- primary school classroom. Objectives/Goals: 1) To illustrate the existence of magnetic force and its function. 2) To classify objects and formulate hypotheses regarding materials and their magnetic properties. 3) To illustrate that magnetic pull is greatest when the object is closer to the magnet; magnetic force passes through objects it pulls, however the magnetic force decreases with distance; a magnet can hold a limited amount of weight.

The activity involves two stages: first children in small groups make predictions and classify various materials such as metals, paper, cork, coins according to their behavior when they come close to magnet. Then, the whole class discusses the categorization of these various materials. Here we focus on the first stage and particularly on one of the children's groups. Our research question is: how the Subject (University Student), used the Tools (materials) and how she shared them to the different groups of children. University student used 4 objects, coins, piece of paper, plastic objects, an aluminum can. She asked children to classify objects according to their magnetic properties. The question was: which materials can be pulled by a magnet and which cannot? The children divided the Tools (materials) in two different categories as follows: A) materials that a magnet can pull, B) materials that a magnet cannot pull. The study of magnets began with an open exploration of magnetic and non-magnetic properties of objects (figure 1). We provided each pair of students with a horseshoe magnet and encouraged them to go on a "magnetic hunt" with their partner. They explored the room, predicting the behavior of different materials when they came close to a magnet.

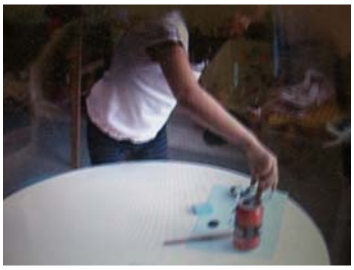


Figure 1: exploration of magnetic properties

University student asked the pupils to present a characteristic for each material and the reason why those materials are pulled by a magnet or not. Results are divided in two categories.

A) Video analysis: From the courses which the students videotaped, we chose a ten minute part, based on the eight-step model (Mwanza 2001). In this stage we notify the children's cognitive obstacles.

1. Children have not a clear opinion about the magnetic materials and the different shaped magnets.
2. Children cannot notice the difference between the magnetic poles.
3. Children believe that the larger a magnet is the bigger magnetic force it applies.

Video studies of our laboratory lessons. We videotaped interactions and analyzed them. The total body of data material includes 3 ½ hours of video material. The goal of our research was to videotape three different groups of students, to evaluate the outcomes with expansive learning and co-configuration (Engeström, 2005, 2007) on one hand, and create new activities on the other.

Evaluation of the videotaped materials as far as magnetic concepts are concerned. In this stage we faced some difficulties, as a great percentage of students were not familiar with the magnetic properties and also with the use of video camera.

Questionnaires and Interviews. The outcomes of the previous stage led us to create a questionnaire and also to conduct an interview with students, so as to learn more about their understanding of the magnetic properties. The questionnaire included thirteen multiple choice questions. The first six questions are content knowledge, the remaining seven questions are procedures of scientific methods. The students were classified according to their performance and then a small group, the best four of them, so as to ensure that we will not face problems of content knowledge, based on cognitive obstacles was selected to continue to the next stage.

Elaboration of data by the use of the Eight- step model (Mwanza, 2001). While working with the four students, they modified the activities and the materials on their own. Furthermore we co-defined the schools (community) where students would practice. The four students were divided in groups of two, they were sent to two different schools for a week. There is a six hour video, which we analysed to reach to certain conclusions.

Students' School Practice. The four students practiced on teaching the activities they had already designed. We video-recorded teaching in order to trace new fruitful activity triangles and evaluate to what extent the outcomes of the previous activities became the objects of some news and so on.

Reflection. A deep discussion with the full community took place and we recorded some conceptual, social and emotional impacts of the whole project.

B) Interviews:

The interviews were taken from the students after the two-week practice in the schools. Students were asked to make some comments and find solutions for the activities planned. Two of three pairs explained us that they faced some problems with the groups of pupils. Many of the children could not attend the other pupils while they worked with magnets.

They felt more confident with the magnetism concept and they noticed the opportunities of adopting alternative methods to achieve their aims. Consequently, all pairs reported that pupils faced some problems with the use of the materials. The most important detail was that this theory encouraged different categories of students to participate in groups (pupils from other countries, pupils with hearing problems).

The results of the study refer to subject-object orientation and especially focus on object. Furthermore, the various factors of the Engeström's triangle - e.g. rules, community, and division of labor – had been studied through analyzing each sub-triangle in relation to the major one (see Figures 2-7 below).

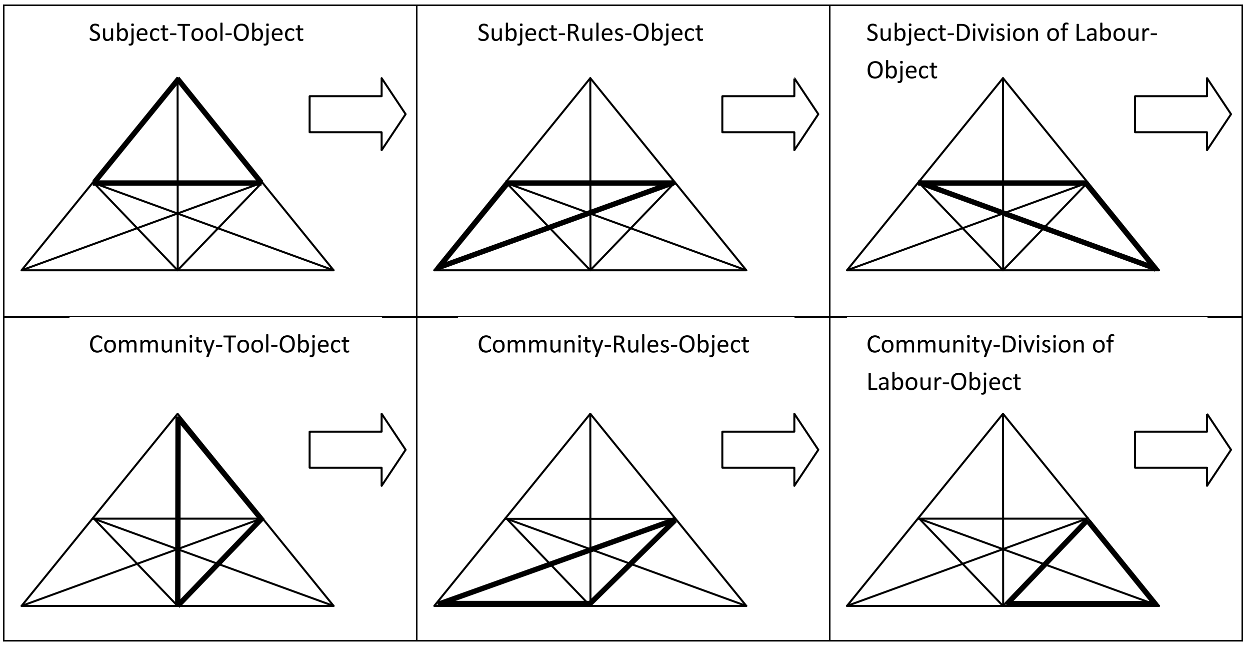


Figure 2-7. Each sub-triangle shows the interaction of 3 factors which produces some different outcomes.

Conclusions and Implications

In conclusion, the results of using cultural-historical activity theory (CHAT) to science education laboratory lessons in the field of Magnetism seems promising. Cultural studies of Science Education seem to be a new potential to make science education a matter for all citizens, or in other words an alternative way to scientific literacy. The theoretical framework of Expansive learning (Engeström, 2005) seems to be appropriate and fruitful for science education researchers. The Mwanza eight- step model seems to be easier to be understood by science education learners. The students' capability to design learning activities could offer criteria to analyze and evaluate learning in university science education laboratories.

In an expansive cycle, decisive actions of individuals within an activity system — which emerge in reaction to and resolution of deep internal contradictions — coalesce to form a qualitatively new mode of joint activity. An expansive cycle can be thought of as the reorchestration of an activity system that occurs when there are shifts within and between its six "corners." (Engeström, 1999: 383):

This case study, even limited, could push forward the boundaries of teachers training in science education. The project was effective and the results came out with the contribution of several groups. At first, school teachers who teach Natural Sciences in the early grades saw this didactical intervention as a chance for improving science teaching in their classes. Also the collaboration with the students at the workshop gave the opportunity to investigate on the way it works in a real classroom situation and discuss their data collection with other students as well as with nursery-school teachers. The most important was that the university students became capable to design activities following the eight steps of Mwanza model (Mwanza, 2001).

References

Cole, M. (1995). Socio-cultural-historical psychology: some general remarks and a proposal for a new kind of cultural-genetic methodology. In: J.V. Wertsch, P.D. Río, and A. Alvarez, (Editors), (1995). Sociocultural studies of mind, Cambridge University Press, New York, p. 187–214.

Engeström, Y. (1987). Learning by expanding: An activity-theoretical approach to developmental research. Helsinki, Orienta-Konsultit.
Engeström, Y. (1999). Activity theory and individual and social transformation. In: Y. Engeström, R. Miettinen, and R.-L. Punamäki, (Editors), (1999). Perspectives on activity theory, Cambridge University Press, New York, p. 19–38.
Engeström, Y. (2005). Developmental work research: Expanding activity theory in practice. Berlin: Lehmanns Media.
Hasu, M, Engeström, Y. (2000). Measurement in action: an activity-theoretical perspective on producer-user interaction, International Journal of Human-Computer Studies, 53 (1), 61-89.
Kuutti, K., (1996). Activity Theory as a Potential Framework for Human-Computer Interaction Research. In B.A. Nardi, (Editor), (1996). Context and Consciousness: Activity Theory and Human- Computer. Cambridge: The MIT Press.
Mwanza, D. (2001). Where Theory meets Practice: A Case for an Activity Theory based Methodology to guide Computer System Design, appear in Proceedings of INTERACT' 2001: Eighth IFIP TC 13 Conference on Human-Computer Interaction Tokyo, Japan.
Nardi, B. (Editor), (1996). Context and consciousness: Activity theory and human-computer interaction, Cambridge: The MIT Press.
Roth, W.M., Tobin, K. (2007). Science, Learning, Identity, Sociocultural and Cultural- Historical Perpectives. SensePublishers, New York, p. 203-204.
Yin, R.K. (1994). Case study research: design and methods. 2nd edition. SAGE Publications, Thousand Oaks, CA.

ATTITUDES OF STUDENTS AND TEACHERS TOWARD GROUP WORK IN TEACHING BIOLOGY

Zaklin Luksa
Gymnasium Cakovec

Diana Garasic
Agency for education, Zagreb

Ines Radanovic
Zagreb University

Abstract

The research was conducted on a sample of 238 high school students and 42 biology teachers in secondary schools in 2007 and 2008. Questionnaires for teachers and students by which their attitudes to group work were examined were used in the research. The students were divided into two groups and questioned before and after the lessons at which the experimental group used inquiry-based learning during group work and the control group mostly used frontal lecture-type classes. The majority of teachers consider themselves qualified for conducting group work, but it is very rarely carried out, very often as pseudogroup work. The change in students' opinions after the experiment can be noticed, but it is statistically insignificant. In spite of the satisfaction with group work that the students show, they still prefer passive teaching methods. Teachers' answers do not correspond to the real behaviour in class. Our results indicate the need of further research to investigate the reasons why group work in biology lessons is used rarely, despite mostly positive attitudes of most teachers to it, as well as the attitudes of the students who practice group work in longer period of time. Therefore they might suggest how to motivate and encourage the teachers to using group work at biology classes more frequently.

Introduction

Group work and cooperative learning are the fields that have been researched very well, with a lot of evidence of positive influence to acquiring learning contents and improving the attitudes related to education and learning contents (Springer *et al.*, 1999; Johnson *et al.*, 2000; Tanner *et al.*, 2003). The more the students are active when learning a topic, the higher is the probability to remember the topic later (Miljevic-Ridicki et al. 2003). A successfully conducted cooperative education influences the development of cognitive strategies, encourages motivation of students and mutual help among the students who are working together (Johnson et al, 1990). Slavin (1995), in the meta-analysis in 1995 pointed out the positive effects of cooperative teaching in primary and secondary school. According to the research he analysed, that kind of a teaching method contributes to higher achievements, positive relationship among students and has psychological advantage if compared with individual and competitive relationships among the students. Is spite of a lot of positive influence, group work is not much liked among teachers (Rotering-Steinberg according to Klippert, 2001). So, the most frequent remarks of both students and teachers are the following: some students do everything, the others do not do anything, everything is done except what should be done; the work within the group is boycotted (Klipert,2001; Pétursdóttir, 2008). Similar attitudes of students and teachers to group work show that problems appear when there is no real interaction among students, and although they sit in a group, the tasks are done individually, with no real cooperation, the reason of which is the fact that group work is not correctly structured in order to ensure group interaction, therefore we cannot talk about efficient experience of learning in a group (Tanner et al., 2003), but the so-called pseudogroup work. In Germany only 8% of teaching time is dedicated to group work, whereas three fourths go to frontal teaching (Klippert, 2001). Research on use and effects of group work in Croatia almost does not exist. The aim of the research was to establish to which extent group work is present in biology teaching, which are the students' attitudes towards this teaching method and to experimentally establish how a well-organised group work influences the change of students' behaviour towards group work.

Rationale

Knowing the present condition and understanding group work should contribute to better understanding of the possibilities of using group work in biology classes. We presume that group work is used very rarely in biology classes and very often it is not well-organised (pseudogroup work), and the majority of teachers do not think they are competent enough for conducting group work in a quality manner. The students feel better when working in groups and it motivates them to further activity, so, therefore they achieve better results, although they need more time to accept the new kind of work which requires more activity of their own. The Aim of the research was: to establish to which extent group work is present in biology teaching, what are the students' attitudes towards this teaching method, and to experimentally establish how a well-organised group work influences the change of students' behaviour towards group work.

Methods

The research was conducted on a random sample of 238 second-grade high-school students and 42 biology teachers in secondary schools in 2007 and 2008. To question the attitudes of teachers, a questionnaire containing nine questions with Likert scale answers (Cohen, Manion, 1980) and two open type questions were used. They were questioned about their attitudes toward group work and how often they practice it in teaching. The students who participate in the research by random sample were divided into two groups, the control one (139 students) and the experimental one (99 students). They were questioned a month before and a month after the lessons when the experimental group used inquiry-based learning during group work, and the control group used mostly frontal lecture-type classes and dialogue. The students' questionnaire asked how much group work compared to other teaching forms influence the success in learning, students' motivation, activity, the way they feel when working in groups, and how much group work (beyond the research) is being conducted in classes. The students questionnaire containing 23 questions with Likert scale answers (Cohen, Manion, 1980) and 3 open type questions were used. Harmonisation of the students' groups was conducted by means of the results of Raven's progressive matrices (Verguts, De Boeck, 2002), the data about gender, the data about age, biology grade and the previous knowledge. Raven's progressive matrices were conducted by a psychologist. The test is non-verbal and therefore excludes the need for any formal school knowledge for its solving and it is sensitive enough for the needs of this research. The previous knowledge was tested by the Institute of Education for the seventh grade of primary school (this is the time when students are taught similar contents as in the second grade of high school).

Results

Comparison the data of the groups (experimental and control) with statistical methods (t-test) about gender, age, biology grade, the previous knowledge, means of the results of Raven's progressive matrices has not shown any statistically significant differences, so the results can be compared (Table 1).

Table 1. Analysis results of Raven's progressive matrices

Group	E (experimental)	C (control)	t-test results (comparison of 2 samples of same variances)	
Number of examinees	93	134	t Stat	**-0,02**
Arithmetic mean	49,44	49,46	P(T<=t) one-tail	0,49
Standard error	0,48	0,40	t Critical one-tail	1,65
Mode	53	52	P(T<=t) two-tail	0,98
Standard deviation	4,67	4,63	t Critical two-tail	1,97
Sample variance	21,77	21,44		

The Results of Students Questionnaire before and after the Experiment

The results of the survey conducted before the experiment show that the students of both groups evaluate the frontal lecture-type teaching as more successful for understanding biology according to the Likert scale (Figure 1.), whereas after the intensive group work the students of the experimental group assess frontal work as less successful although this difference is not statistically significant.

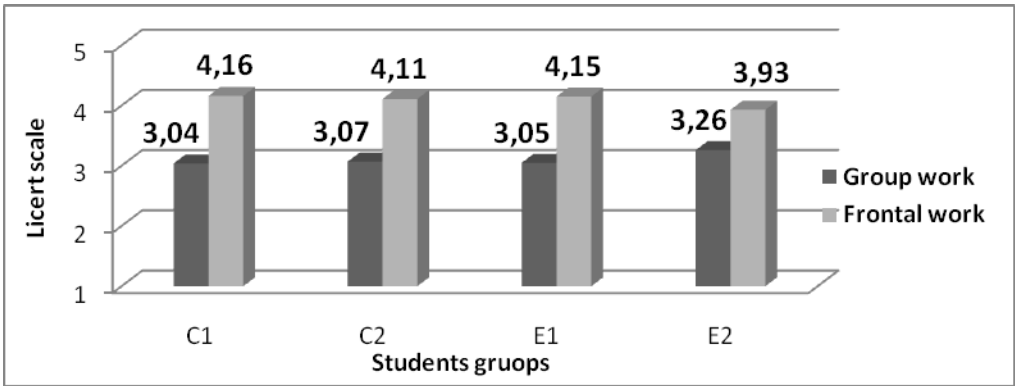


Figure 1. The Results control (C) and experimental (E) group of Students Questionnaire before (C1,E1) and after the Experiment (C2, E2) on question : Group work / Frontal work …helps me understand the content of biology.

The results of all groups before and after the experiment showed that students considered group work as being fun, with more freedom, the students are more engaged in cooperating with each other (Figure 2.).

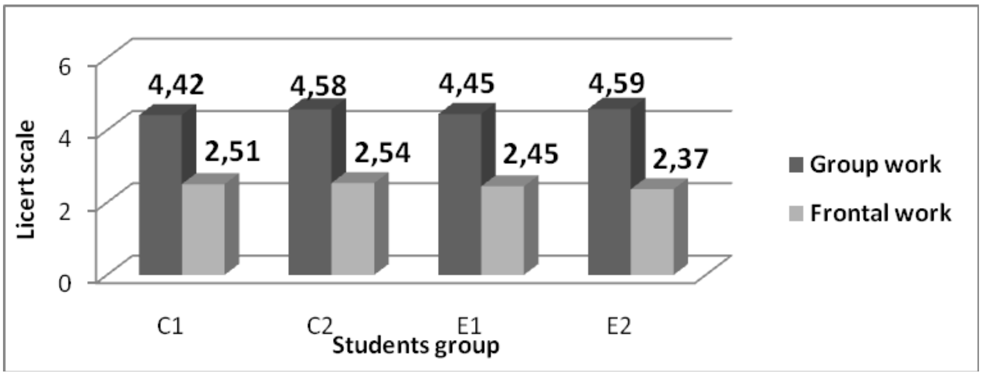


Figure 2. The Results control (C) and experimental (E) group of Students Questionnaire before (C1,E1) and after (C2, E2) the Experiment on question - Group work / Frontal work … provides more fun when learning.

The results of all groups before and after the experiment indicate that students evaluate that group work should have greater involvement of students in the classroom (Figure 3.).

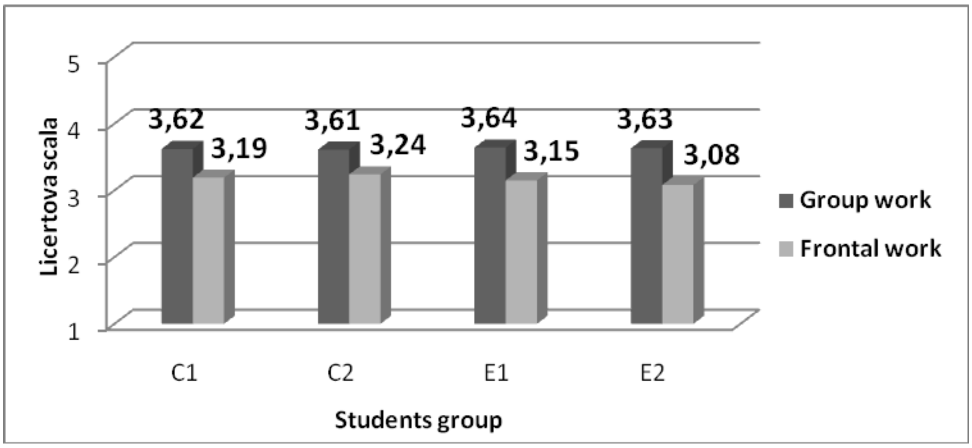


Figure 3. The Results control (C) and experimental (E) group of Students Questionnaire before (C1,E1) and after (C2, E2) the Experiment on question - Group work / Frontal work …required greater activity and involvement of students in the classroom.

All students consider group work more amusing, freer, they are more engaged in it and they get to know each other better, and although they say that active participation in class contributes to better acquisition of the contents, it is contradictory to the opinion that they consider frontal work better for more successful learning (Figure 4.). This is also noticed from the students' statements that during group work their thoughts wander around and they disturb each other more than during other teaching and learning activities. In spite of the fact that more than half of the students say that they feel more comfortable during group work, they are more inclined to those teaching methods that do not require too much effort. This indicates to non-existence of habits of learning through work and shortage of these teaching methods in class. These results indicate that group work is often not well organized and that there is no real cooperation between students. These results speak in favour of the hypothesis that teachers are not sufficiently trained to carry out group work.

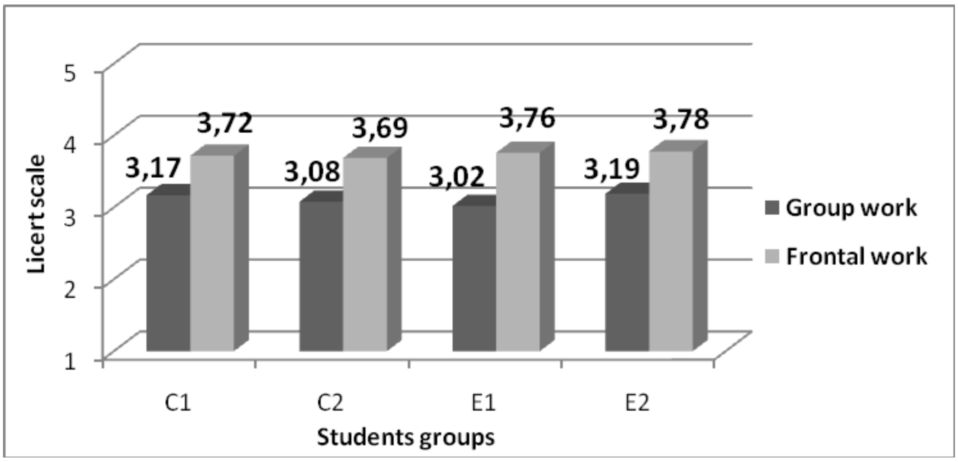


Figure 4. The Results control (C) and experimental (E) group of Students Questionnaire before (C1,E1) and after (C2, E2) the Experiment on question - Group work / Frontal work … allows successful learning.

The Results of Teachers Questionnaire

While the results of a survey of students to speak in favour of the thesis of weak organized and carried out group work results of a survey of teachers in conflict with it. Thus, even 88.13% of teachers considered to be good, very good and very well educated for the implementation of group work (Figure 5.).

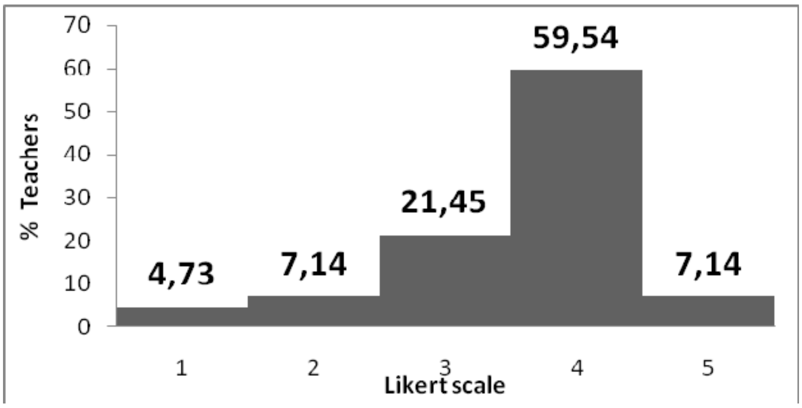


Figure 5. Results of self-evaluation of teachers regarding the implementation of training for group work in teaching biology (I am educated enough to conduct group work in teaching biology.).

It is unusual that on the other hand up to 42.3% of teachers think that the students in this way acquire higher-quality knowledge than when frontal approach is used, 59.3% think that their students are more content with group work than with frontal work, 38% think that the students use time more efficiently during the lesson when group work is used.

The results of the survey conducted among teachers show as that only 21.4% of teachers more often or regularly use group work in biology classes, which is comparable to Klipert's results (2001.). The results are supported by the students' opinion that group work is very rarely used.

As disadvantages the teachers more frequently mention the following: it is not economical and it wastes time, preparations for high-quality group work are too demanding, chaos in groups, nonparticipation of all, i.e. participation of only some individuals, avoiding responsibilities, problems with organisation because a double period is needed as well as the lack of material and equipment for all groups and students. In addition, classes are too big for this kind of work, students do not take it seriously because they experience this kind of work as a game, weak feedback and difficulties with marking students. The mentioned shortcomings correspond to the shortcomings which Klarin (1998) and Klipert (2001) state in their works.

Comparison of results of surveys of students and teachers

The results show significant differences in the opinion of students and teachers about what form of students achieve better results in two questions. In question 3. (Table 6. Students achieve better learning results with frontal work) students agree with the statement, but the teachers think the opposite. In question 4. (Table 6. - Students are active in the classroom when doing group work.) teachers agree with the statement, but the students think the opposite.

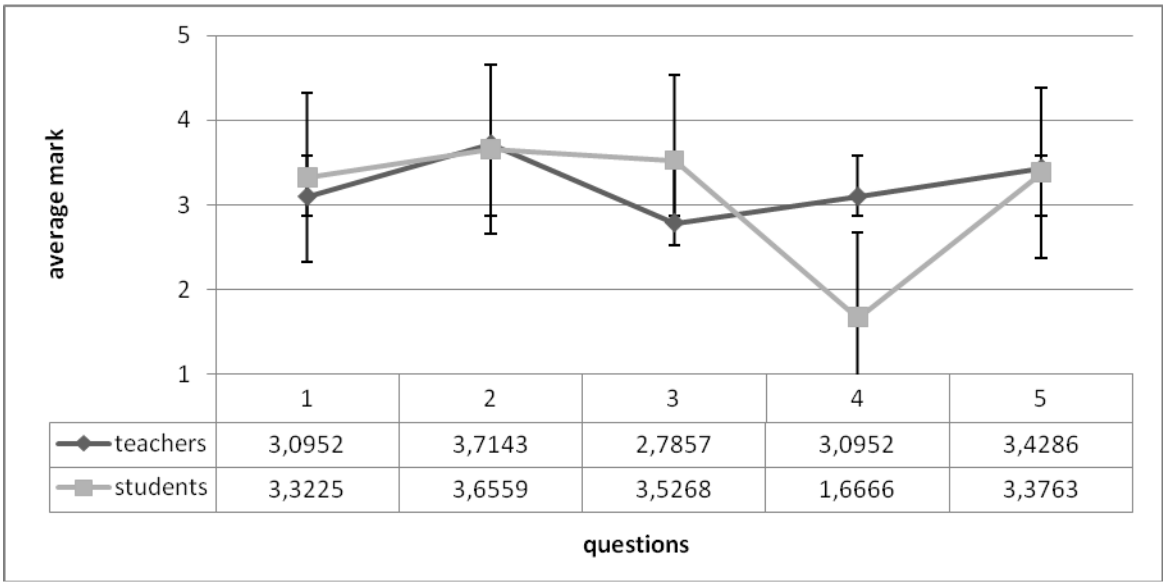


Figure 6. Comparison of responses of students and teachers the same questions from the survey.

Conclusions and Implications

In biology teaching in Croatia group work is practiced rarely. Statistically significant change of the students' attitudes toward group work after having conducted intensive group work classes has not been proved, in spite of statistically better testing results compared to the students who were taught frontally (Luksa, 2007). The probable reason is that the new teaching method was applied for a too short period of time and included active teaching methods for older students. It is shown that 88% of teachers think they are trained enough to conduct group work and consider that enough high-quality knowledge can be acquired by using group work in class. Despite that, the teachers are not too much inclined towards using group work in biology teaching. Students also estimate that group work is rarely used in biology classes. In favour of the fact that teachers conduct group work only formally, without any interaction of students, speaks the fact that in spite of the satisfaction with group work they have shown, the students still prefer the teaching methods in which teachers have the leading role whereas they are only passive spectators or they, occasionally, participate in discussions. These results show that it is likely that the teachers answered the questions in the way they thought it was expected from them, whereas in a class they behave totally different. This is the reason why the students are not well prepared and trained for group work, so the change in attitudes of students after a month-long experiment is noticeable but not statistically significant. Our results point to the need for further research with the aim of more detailed investigation of the reasons why group work in biology classes is not used more often in spite of generally positive attitudes of the majority of teachers about it, and the attitudes of students who encounter this kind of teaching in longer period of their education. In this way we could suggest how to motivate and encourage teachers to use group work in biology classes more frequently.

This research was part of scientific project “Competences of pupils attending science and biology classes” No. 119-0091361-1223, supported by the Ministry of Sciences, Education and Sports of the Republic of Croatia.

References

Cohen,L., Manion, L. (2002). Research methods in education, London, Routledge, 1980.Verguts, T., De Boeck, P., The induction of solution rules in Raven's Progressive Matrices test, European Journal of Cognitive Psychology, 14 (4), 521-547

Johnson, D. W., Johnson, R. T. (1990). Cooperative Learning and Achivement. Shlomo Sharan (Ed.). New York, 22-37.

Johnson, D.W., Johnson, R.T., and Stanne, M.E. (2000). *Cooperative Learning Methods: A Meta-Analysis*. Cooperative Learning Center website, Retrieved from http://www.clcrc.com

Klippert, H. (2001). How to successfully learn in a team. Collection of practical examples. EDUCA. Zagreb.

Luksa, Z. (2007). Academic achievements of students by using group work in biology lessons. Napredak. Juornal for pedagogic theory and practice, vol.148/No. 4, 549-564.

Miljevic-Ridicki R., Miljkovic D., Pavlicevic-Franic D., Rijavec M., Vizek-Vidovic V$., Vlahovic-Stefic V.,Zarevski P. (2003). Teachers for teachers: examples of the implementation of principles of Active/efficient school, IEP, Zagreb,5-10.

Pétursdóttir B. (2008). What matters is communication. Secondary students' responses toward a science learning environment, NFSUN, Reykjavik, Retrieved from http://www.symposium9.khi.is/synopsis/nfsun9_submission_113.doc

Slavin, R. E. (1995). Reasearch on cooperative learning and achievement: What we know, what we need to know, Contemporary Educational Psychology, 21, 43-69, Retrieved from http://www.successforall.net/research/

Springer, L., Stanne, M.E., and Donovan, S. (1999). Measuring the success of small-group learning in college level SMET teaching: a meta-analysis. *Review of Educational Research* ,69, 21–51.

Tanner, K., Chatman, L. S., Allen, D. (2003). Approaches to Cell Biology Teaching: Cooperative Learning int he Science Classrom – Beyond Students Working in Groups. Cell Biology Education, vol. 2, 1-5.

GLOBAL WARMING: CAN IT EMPTY OUR RICE BOWL?

Pongprapan Pongsophon
Kasetsart Univeristy

Saouma B. BouJaoude
American University of Beirut

Naruemon Yutakom
Kasetsart University

Abstract

The purpose of this study was to investigate Thai students' understanding of global warming. One hundred and six Grade 7 students filled out an open-ended questionnaire about global warming. In the first part of the questionnaire, selected key words related to global warming were listed asking the students if they had ever heard about the terms and if so, how they would define them. In the second part, their ideas were probed further. The topics investigated included the causes, processes, and consequences of global warming. Results showed that the majority of students had heard only the keywords splashed in mass media and their given definitions reflected alternative conceptions. The majority of students held partial understanding of the causes of global warming. For example, they thought that only CO_2 and CFCs were greenhouse gases. They held specific misconception about the mechanism of global warming. They thought that ozone layer depletion caused global warming and atmospheric greenhouse worked in the same way as an actual greenhouse did. Moreover, they held partial understandings of the consequences and cures of global warming. The consequences were perceived as being large-scale, human-centered and catastrophic. The cures were commonly viewed from individual rather than multi-dimensional perspectives. There seemed to be a general conflation of ideas whereby many pro-environmental actions that are not connected with global warming were seen as helping to reduce it. Some implications for promoting students' scientific literacy on this issue were discussed.

Introduction

Thailand is a country in the midst of change. Since 1980, it has undergone considerable industrial development and urbanization. With high economic growth rate, it has faced environmental deterioration stemming from large increases in energy consumption. In 2004, Thailand's total energy consumption was 0.7 percent of total world energy consumption. Thailand is a net importer of oil and natural gas. In 2005 daily oil, natural gas, coal consumption exceeded domestic daily production. This results in the increase in greenhouse gas emissions. CO2 emissions were projected to increase from 282 Tg (Teragrams) in 2010 to 475 Tg in 2020; methane (CH4) emissions from 2.1 Tg in 1994 to 2.20 Tg by 2010, and 2.24 Tg by 2020 (Patanavanich, 1991). The consequences of global warming have been continuously reported. The northern areas of Thailand were hit by flash floods and unexpected weather conditions. The sea level of Andaman Sea rose. The wind was so strong and unpredictable that small fishing boats could not venture into the sea. The decline in annual flow of the Mekong River has contributed in increasing water stress and declined agricultural production. Economically, in particular, Thailand is one of the world's largest rice exporters. As the world gets hotter, floods, droughts and rising sea levels could push Thailand's rice yields down significantly (McDonald-Gibson, 2007). The animals and small microbes that give the soil its nutrients are very sensitive to heat and humidity. Flood and heavy rain would erode the soil and destroy the nutrients.

Global warming has become a hotly debated issue and a big concern of the nation. Thai media has sensitized and informed the public about the global warming issue. A poll conducted recently among 2,191 people aged above 18 in Bangkok and surrounding areas by Assumption University in September 2007 showed that global warming worried 97% of all respondents (Thai News Agency, 2007). To accomplish sustainable development, Thailand must establish goals leading to the reduction of greenhouse gas emissions. Thailand signed the United Nations

Framework Convention on Climate Change (UNFCC) in 1992 and the Kyoto Protocol in 1999 (IPCC, 2001). Consequently, Thailand is preparing measures to reduce greenhouse gas emissions including regulations, fiscal incentives, information, and research, development and demonstration. In response to this issue, the Thai Ministry of Education with support from the private sectors ran a series of campaigns on global warming in schools nationwide. They have sought to teach the fundamentals of global warming and the complexity of the issue to a school audience. Government concern over the environmental issues ensured their place in the National Curriculum Framework of Thailand, which emphasizes that school science should inform students about current environmental problems and empower them to make positive contributions to this issue (IPST, 2002). Teachers in general and science teachers more particularly can support these campaigns because they are in direct contact with students. However, for teachers to succeed in communicating ideas about global warming, they need to understand student's existing ideas about the concept, an understanding that will help them to identify appropriate learning experiences to help students change their alternative conceptions and encourage them to take action on the issue. There was a considerable volume of literature on students' alternative conception on global warming (Dove, 1996; Fisher,1998; Kilinç et al., 2008; Mason & Santi, 1998; Stamp, et al., 2007) but less research had been conducted with students in the Asia-Pacific region. Consequently, the purpose of this study was to investigate Thai students' understanding of global warming in the hope that the results would help science teachers to "start from where students are" when they design instructional materials.

Research Question

How Thai lower secondary students explain causes, process, consequences and cure of global warming?

Theoretical Framework

Students' ideas in science and about global warming

Children have developed ideas about the natural world around them before coming into a classroom. They built ideas through sensory experience while interacting with environment (Driver, 1994). They tried to make sense of the natural world around them. This explanation was meaningful for them. It is an everyday way of talking and thinking about that phenomenon. It is logical, coherent and internally consistent even may differ substantially from the scientific view. Children's ideas were, therefore, robust and persistent into adulthood despite formal teaching. Students have difficulties to integrate any new information with their cognitive structures since the existing knowledge is integrated and resistance with their experience. Children struggled to come across scientific ideas which often seemed to go against common sense. Acceptance of scientific ideas does not lead automatically to elimination of misconceptions about the same issue. In many cases, children hold more than one explanatory model and bring up one rather than one another in a particular context. Previous studies reveal commonality in the conceptual models that children from different countries and backgrounds are using. Their ideas are not idiosyncratic, nor are they heavily culturally dependent. Previous studies suggested that student's ideas about global warming from different countries share commonality (Dove, 1996; Fisher,1998; Kilinç et al., 2008; Mason & Santi, 1998; Stamp, et al., 2007). For example, students appeared to confuse the causes and mechanism of global warming with those of ozone layer depletion, another major environmental problem. They explained that ozone holes allowed solar radiation to penetrate. Students misunderstood about the consequences of global warming. They associated global warming with an increased incidence of earthquakes, diseases such as cardiac problems and skin cancer, fish poisoning, unsafe drinking water and food poisoning.

Scientific Background of Global Warming

This section was the review of literature about global warming (Houghton, 1997; Hardy, 2003). It was used as a reference in developing an assessment tool and analyzing and interpreting the data.

Causes of Global Warming

The sun's energy warms the Earth's surface and its atmosphere. As this energy radiates back toward space as heat, a portion is absorbed by heat-trapping gases in the atmosphere. This keeps the Earth in a temperature range that allows life to flourish. These gases are known greenhouse gases. There are a number of green house gases such as carbon dioxide (CO_2), Nitrous Oxide (N_20), methane (CH_4), water vapor, sulfur hexafluoride, and CFCs, etc. In order, Earth's most abundant greenhouse gases are water vapor, carbon dioxide, methane, nitrous oxide, ozone, and

CFCs. When these gases are ranked by their contribution to the greenhouse effect, the most important are: water vapor, which contributes 36–70%, carbon dioxide, which contributes 9–26%, methane, which contributes 4–9%, and ozone, which contributes 3–7%. The water vapor, most prevalent powerful greenhouse gas holding onto 2/3 of the heat trapped by all greenhouse gases, is a consequence of global warming. As the Earth heats up relative humidity increases, allowing the planet's atmosphere to hold more water vapor, causing even more warming, thus a positive feedback scenario. Atmospheric CO_2 is produced by a number of sources including burning of fossil fuels by cars, electricity generating power plants, airplanes; and deforestation. Methane, is derived from rice paddies, bovine flatulence, bacteria in bogs, and fossil fuel production. Nitrous oxide (N_2O) or laughing gas is produced either naturally in ocean forest or by humans in nylon and nitric acid production, the use of fertilizers in agriculture, and cars with catalytic converters.

Process of Global Warming: Greenhouse Effect

The Earth receives energy from the sun in the form of visible light. Most of this energy is not absorbed by the atmosphere since it is transparent to visible light. The energy hits and warms up the earth. It then radiates back to the atmosphere in infrared range. The energy is, this time, trapped since the greenhouse gases are not transparent to infrared. They absorb thermal infrared radiation. As a result of its warmth, the atmosphere also radiates thermal infrared downwards to the Earth's surface. This keeps the Earth's temperature steady; not rapidly heated and cooled. This mechanism is fundamentally different from the mechanism of an actual greenhouse which isolates air inside the structure so that heat is not lost by convection and conduction. Anthropogenic global warming is a global mean temperature anomaly trend that results from an enhanced greenhouse effect mainly due to human-produced increased concentrations of greenhouse gases in the atmosphere.

Consequences of Global Warming

Increasing global temperatures are causing a broad range of changes. For example, sea levels are rising, and land ice in the poles is melting , and the amounts and patterns of precipitation are changing. Moreover, there is an increase in the frequency, duration, and intensity of extreme weather events, such as floods, droughts, heat waves, and tornadoes. Other effects of global warming include lower agricultural yields, further glacial retreat and disappearance, reduced summer stream flows, and species extinctions. In addition, diseases like malaria are returning into areas where they had been eliminated earlier.

Methods

Participants and Setting

The respondents were 106 Grade 7 (Year 12-13) students; 49 males and 57 females. Eighty nine percent of students are 17-18 years of age. They were from two classes which had been randomly sampled from ten mixed ability classes. The majority of the students came from middle and upper class family. Ninety-eight percent of students were Buddhist. The rest encompassed Christian and Muslim. Sixty percent of them lived in the district where school was located; a central district of Bangkok. This district was known as the one facing the serious massive traffic jams. The major reason is the continued popularity of private automobiles, and extensive consumer credit for automobile purchases. Another reason is the ongoing construction of mega-project; expressways, tunnels and flyovers, sky trains (BTS) and underground metro systems, and Bus Rapid Transit. The school was a special large size public, co-educational secondary school with more than 4,000 students. The school facilities were fully equipped and still in very good condition. The school was all given good and continuous financial support by parents' association. The school allocated this fund for the benefit of education, welfare and a healthy community within the school. The school has science self-learning center where the students had access to numerous science textbooks, magazines, science fiction, posters, and internet. The center was, in addition, equipped with multimedia such as a video player, televisions, and a projector. To create permanent public awareness about energy conservation, the school implemented Green Learning Room program under financial support by Electricity Generating Authority of Thailand (EGAT). The green learning room offers a variety of interactive learning tools to educate school children on the link between energy use and environment.

Instrument

An open-ended questionnaire about global warming was used to collect data in this study. The questionnaire was validated by content experts and piloted with a cohort of 30 junior high school students similar to the target group. It took two hours to complete. The questionnaire consisted of two parts. The first part was a list of nine keywords related to global warming asking students if they had ever heard of such key words and if so, from where and how they would draw definitions. The keywords can be divided into three groups; causes, consequences, and cures of global warming. They were recommended by the content experts since they were fundamental concepts of global warming that everyone should know as scientifically literate persons. They were causes [1. greenhouse effect]; consequences [2. climate change 3. coral reef bleaching]; and cures [4. Kyoto Protocol 5. biodiesel 6. gasohol E20 7. diesel B5 8. non-renewable energy 9. renewable energy]. In the second part, their initial ideas were investigated in-depth. In each question, the students explained their ideas and were allowed to draw a picture or a concept map to illustrate those ideas if needed.

Data Analysis

The researcher employed quantitative and qualitative data analysis to make sense of the data. In first part of the questionnaire, the percentages of the heard and not heard were calculated. The definitions of keywords given by the heard were examined. It was classified into scientifically or non-scientifically consistent. The second part, students' explanation of all students were gathered and thoroughly read. Codes were identified, refined and revised until a category system emerged. The category was examined its correction against scientific conception and then classified into one of the five levels of conceptual understanding in a conceptual continuum developed by Abraham and Wilkinson (1994); Sound Understanding (SU), Partial Understanding (PU), Partial Understanding with Specific Misconception (PU&SM), Specific Misconceptions (SM), No Understanding (N). Each level is described as follows:

- Sound Understanding (SU): Responses that included all components of the validated response;
- Partial Understanding (PU): Responses that included at least one of the components of validated response, but not all the components;
- Partial Understanding with Specific Misconception (PU&SM): Responses that showed understanding of the concept, but also made a statement, which demonstrated a misunderstanding;
- Specific Misconceptions (SM): Responses that included illogical or incorrect information;
- No Understanding (N): Repeated the question; contained irrelevant information or an unclear response; left the response blank.

Assigning conceptions to levels of understanding was validated by two independent coders. The percentage of students holding ideas of a certain conceptual level was calculated and compared with those of other levels. This showed the profile of students' conceptual understanding on global warming concepts.

Results and Discussion

Results are two sections. In the first section, we present students' awareness of the keywords about global warming. This is followed by second section; Students' Conceptual Understanding of Global Warming which addresses causes, consequences, and consequences of global warming.

Students' awareness of keywords about Global Warming

The table 1 shows the percentage of students who had heard and never heard the nine keywords about global warming. For the heard, the percentages of those having scientific and non-scientific conception were presented. Ninety seven percent of students had heard about greenhouse effect. However, most of them held alternative conceptions about this concept. Some students held partial understanding; for example, they thought CO2 was the only greenhouse gas; and, not in details, the presence of this gas led to an increase in temperature of the Earth. As for the consequence, seventy five percent of students had heard about climate change but all of them held alternative conception. They thought that it was a fluctuation in whether condition within a short period of time in certain region unexpectedly. Scientifically, climate change is any long-term change in the statistics of weather over durations ranging from decades to millions of years. It can be manifest in changes to averages, extremes and may occur in a specific region or for the Earth as a whole. Eighty one percent of the students have never heard about coral bleaching. For those who had heard their views yet were not aligned with scientific conception. Some of them though due to water pollutions or increase in ocean temperature, the coral loss its natural color. They did not understand that the color is of zooxanthealle, photosynthesizing, unicellular algae that have a symbiotic relationship with the hosting coral. Under stress, the corals expel their zooxantheallae, which leads to a lighter or completely white appearance.

As for keywords related to the cures of global warming, ninety one percent of students had never head of Kyoto Protocol. The rest consisted right and wrong answers by the similar percentage. The Kyoto Protocol is a protocol to the United Nations Framework Convention on Climate Change (UNFCCC or FCCC) (IPCC, 2001), an international environmental treaty with the goal of achieving "stabilization of greenhouse gas concentrations in the atmosphere at a level that would prevent dangerous anthropogenic interference with the climate system." As for the alternative energy, ninety one and seventy seven percent of students had heard of biodiesel B5 and gasohol (E20) respectively. However, more than forty nine percent of students could not identify about their raw materials and the production process. Some of the students had got confusion between the two. As for biodiesel B5, however, the percentage of those had ever heard was dropped to forty one and almost of which held alternative conception. As for renewable energy and non-renewable energy, the percentages of students having heard and not having heard of the terms were quite the same around fifty percent. The majority of the heard, however, held alternative conception.

It is obvious that the majority of Thai students from this study had ever heard of biodiesel and gasohol. This might be the result of serious and continuous campaign through mass media by the Thai government. The students mentioned to mass media, school or direct experience as most frequent source of information about global warming and alternative energy. This finding was in accordance with the study conducted by Kilinç and colleges in 2008 who discovered that school and television was the most popular source. Newspapers and the internet played less of a role in providing information about this issue. However, Kilinç noted that the issue of global warming presented in the media was quite superficial and alternative to scientific conception. This misled the public. The term "global warming" is self-explanatory, and it is easy to imagine how students would make a mental link between deserts and high temperatures. The Ministry of Energy of Thailand has put into vigorous practice and assures that the country is prepared to forge ahead the use of gasohol, biodiesel and NGV (Natural Gas Vehicle) as primary alternative energy and to promote clean energy together with environmental protection. Thailand has set up serious efforts to reduce oil imports by replacing at least 20% of its fossil fuel consumption with renewable energy sources such as ethanol and biodiesel. Biofuels are also seen by the government as an opportunity for rural development and trade (Patanavanich, 1991).

Alternative energy stations are now found every corner throuout Thailand. Thailand currently sells gasohol (commercially known as E10 and E20) and B2 and B5 (2, 5% biodiesel with 98, 95% petroleum diesel respectively) through its service stations of the state-owned companies PTT and Bangchak. There were 3,822 gasohol service stations in Thailand 2007. In 2008, 40 stations in Greater Bangkok sell E20. B2 is available at all stations throughout Thailand; 976 stations offer B5 in Greater Bangkok. E20 compatible vehicles are available in Thailand from Ford, Toyota, Honda, and Nissan. (Patanavanich, 1991).

Students' Conceptual Understanding of Global Warming

Causes of global warming

The majority of students (49 %) held partial understanding about greenhouse gases. They thought that there were only a few types of greenhouse gases and they could not identify all their sources. "CO2 is the cause of global warming. It is emitted from factories and cars during traffic congestion" (M024: Male students#24). Another girl said (F074), "CFCs and CO2 are the principle greenhouse gases. They both come from human activities. CFCs are in a long thin tube at the back of an older refrigerator. We used this chemical, when it turns to a vapor, to soak up some of the heat in the freezer compartment. This gas was found harmful to the environment if it leaks out". Greenhouse gases such as methane were hardly mentioned and nobody referred to nitrous oxide or water vapor. Greenhouse gases, in fact, consists of varying amount of water vapor and cloud (97%) with the remainder being gases like CO2, CH4, Ozone, and N20 (De Freitas, 2002). Media coverage has focused on carbon dioxide and therefore it was not surprising to find wide realization among the students that CO2 was an abundant greenhouse gas. This was consistent with previous studies such as Mason & Santi (1998), Stamp, et al. (2007), Kilinç, et al. (2008) that of carbon dioxide and CFCs were well known as a greenhouse gas by eighty percent of the respondents. A few proportion of students mention methane and nitrogen dioxides. Fewer of the students realized that tropospheric ozone could act as a greenhouse gas. In addition, the students thought that all sources of greenhouse gas emission were involved with human activities such as burning fossils fuels and deforestation. In fact, this gas as well as other greenhouse gas occurs naturally in the atmosphere.

Table 1 : Students' awareness of the keywords related to Global Warming Issue.

Keywords	Awareness		Student's definition	
	Never heard (%)	Heard (%)	Scientific conception (%)	*Alternative conception (%)
1.Cause				
1.1 Greenhouse effect	3.77	96.23	1.88	94.35
2.Consequences				
2.1 Climate change	25.47	74.53	0	74.53
2.2 Coral reef bleaching	81.13	18.87	0	18.87
3.Cures				
3.1 Kyoto Protocol	91.51	8.49	4.72	3.77
3.2 Biodiesel	10.38	89.62	32.07	57.55
3.3 Gasohol E20	23.58	76.42	27.36	49.06
3.4 Diesel B5	59.43	40.57	4.72	35.85
3.5 Non-renewable energy	50.94	49.06	16.04	33.02
3.6 Renewable energy	48.11	51.89	11.32	40.57

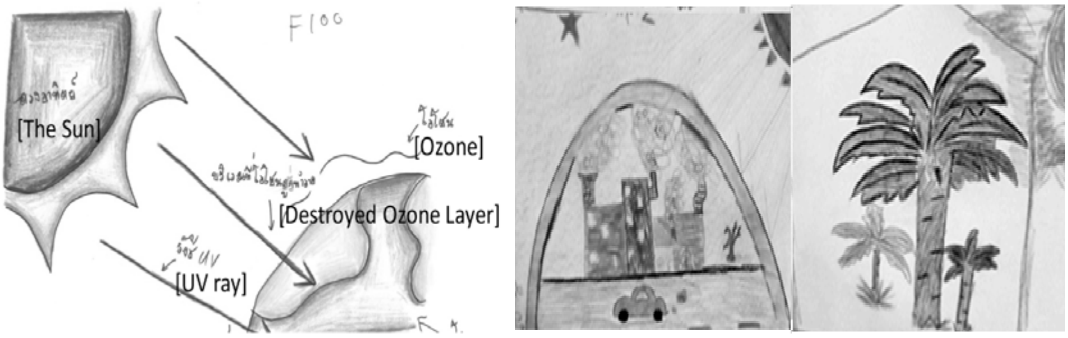
* Non-scientific based definitions are the combination of PU, PU/SM, SM and NU.

Thirty three percent of students held partial understanding with specific misconception (PU & SM). Seven percent of students held specific misconceptions (SM). For example, these students thought that carbon monoxide, particulate matter such as dust and diesel exhaust, and extra sunlight from the sun were greenhouse gases. Noticeably, some students viewed CO2 as a pollutant. In fact, CO2 is essential to life on earth. It is required in the process of photosynthesis for plant growth and reproduction. Greenhouse gases are responsible for keeping the Earth warm; without them, the Earth would be frozen and lifeless. Answer of a number of students could not be categorized because they provided general statements such as "greenhouses are toxic gases in the air"

Process of global warming

Approximately thirty four of the students held specific misconceptions (SM) about the mechanism of global warming. They thought that global warming was a result of ozone depletion. In her drawing (F100, Fig. 1a), there was a big hole in an upper layer of atmosphere (stratospheric ozone), so the sunlight could penetrate and warm the Earth's surface. This reveals a conflation in the minds of students of these two major environmental issues because they appear to perfectly match in the minds of students. Ozone depletion sounds to be a logical cause of global warming because the extra sunrays come through holes in the ozone layer and make the earth hotter. This confusion was consistent with the findings of the previous studies such Dove, 1996; Fisher,1998; Kilinç et al., 2008; Mason & Santi, 1998; Stamp, et al., 2007. Francis et al. (1993) explained this that Alternative conceptions might be the fusion of ideas in the children's conceptual structures". They may know, mainly by information coming from mass media, that there are two major environmental problems such as global warming and ozone layer depletion, but they cannot fully understand and distinguish the underlying different mechanism, the conceptual complexity and linkage involved.

Forty six percent of the students held partial understandings with specific misconception (PU&SM). These students could correctly specify greenhouse gases and identify their sources, yet they confused atmospheric greenhouse with actual greenhouses that are used in agriculture. Many of these students drew a thick layer of CO2 which sun ray could penetrate (Fig. 1b, c). The sun rays then heated up the ground. The leftover reflected and rose up to the atmosphere but could not go beyond the atmosphere. It was retained underneath the layer just like a greenhouse. In fact, a greenhouse in agriculture and atmospheric greenhouse operate differently, the former works primarily by preventing convection; the latter, however, reduces radiation loss. Source of confusion might come from "heat trapping" analogy of how a greenhouse limits convection to how the atmosphere performs a similar function but through the different mechanism of infrared absorbing gases. In addition, a few students could distinguish incoming, solar, short wave radiation from outgoing, long-wave, trapped rays. Most students thought it was solar, not terrestrial radiation, which was being emitted from the Earth's surface and becoming trapped by greenhouse gases. The absorption of gases being dependent on wavelength was largely unknown for the majority of students.



(a) (b) (c)

Figure 1. (a) A drawing of F100 to show ozone depletion as the mechanism of global warming (b, c) Drawings show greenhouse effect working as a greenhouse in agriculture

Consequences of global warming

Eighty seven percent of students held partial understanding of the consequences of global warming. Their responses focused on large regional or global climatic change, unexpected natural disasters and human health issues such as drought, flooding, heat waves, and declining perennial polar ice cap in Arctic. These views might have been influenced by media in which such catastrophes have been in the spotlight (Corbett & Durfee, 2004). Noticeably, Tsunami hit the western coastline of Thailand was exemplified by over 23 percent of the responses. A girl thought, "Global warming brings us a great loss. The effects of Tsunami in Phuket or those of Hurricane Katrina in New Orleans are devastating and costliest" (F021). These students, however, did not elaborate on how these physical changes threaten the economy, society, and ecosystem. A few percent of students were concerned about a substantial reduction in crop yields that could lead socio-economic crises. The majority could have not related rising temperature, longer droughts, flood plain, to rain-fed agriculture. None of them saw ecological effect of global warming. They did not see how increasing global temperature affected physical and biological factors in an ecosystem and how this would trigger ecological change such as changes in range and seasonal behavior in certain species of plants and animals, reduction in ecological productivity and survival rate.

In short, most of their views were large-scale, human-centered, sudden and catastrophic rather than being diverse and gradual. However, a body of observations from Third Assessment Report by IPCC reveals a collective, gradual feature of the consequences of global warming and climate system. Examples of observed climatic changes are the increase in global average surface temperature of about 1 °F in the 20th century; the decrease of snow cover and sea ice extent and the retreat of mountain glaciers in the latter half of the 20th century; lengthening of the growing season in middle and high latitudes etc. In part, this may be due the descriptive terminology "global warming" and the use of logical thinking to derive their answers. Global warming describes the Earth as getting hotter, thus students might have thought that an intuitive consequence could be polar ice caps melting, a cause of sea level rise and flooding. The students, in addition, might be well informed about the physical consequences of global warming from strong media images. A number of misconceptions were found in a few students. They thought, for example, the global warming would result in an increase in the prevalence of skin cancer. This can be result of misconception about ozone depletion as the mechanism of global warming. Some students also thought that there would not be more rainfall. The main reason being that was a warmer climate would, by implication, also be drier. Some students thought that global warming resulted in higher sunshine level. This stimulated photosynthesis.

The cures of global warming

The Ninety two percent of students held partial understanding to the cures of global warming. Most students thought measure for reducing global warming could simply be undertaken at the level of individuals. They suggested a number of ways to cut down the level of carbon dioxide in the atmosphere such as planting more trees, recycling paper, saving electricity, riding a bicycle etc. However, a small number of students discussed the role of government, the private sector, state agencies, local authorities, and other related parties in solving global warming. They did not realize that some of these measures required political action. One powerful way that was entirely dismissed in student responses was taking the opportunity, as well informed individuals, to participate in public hearings, conversations, and forums on global warming issues or the like in order to establish programs and policies that are good for the environment. Seven percent of students held partial understanding and specific misconception. These students misunderstood about the cures of global warming. Many helpful pro-environmental actions in their list did not contribute to a solution of global warming. For instance, they thought draining the street more often, using unleaded petrol and using air filter in the house would reduce global warming.

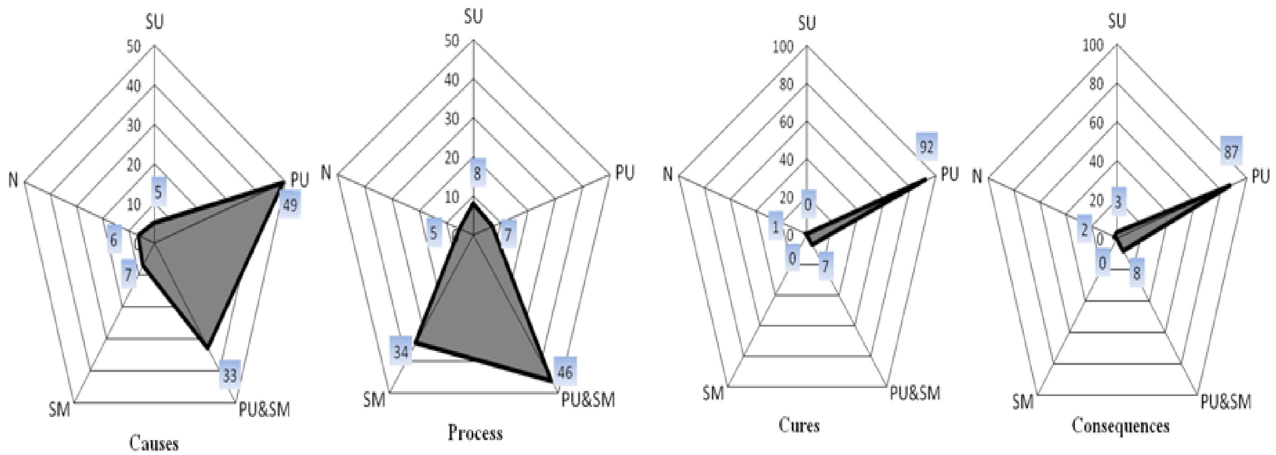


Figure 3. The profiles of conceptual understanding of global warming concepts

Conclusion and Educational Implications

Finding of this study can provide teachers with useful information that can be used to develop relevant, meaningful, and effective instructional materials to help student change the alternative conceptions. A number of resources for teaching climate change are now available and are accessible easily online. These website offer comprehensive information on the science of global climate change, potential impacts, adaptation and mitigation strategies, and policies; detailed lesson plans; teaching materials and experiment. The well-designed laboratory exercises could demonstrate the causes, mechanism and consequences of global warming together with a classroom discussion can provide teachers with situations that cause students to engage, explore and extend their knowledge. In the exploration, the students would experience cognitive conflicts, results in a fruitful breeding ground for knowledge revision and more robust understandings of global warming concepts. Global warming is a great topic for students to study because it integrates so many subjects: energy, environment, geography, politics, chemistry, biology, economics, and more. The teachers, therefore, encourage students to use analytical thinking and to exercise their abilities to research, think and understand complex issues. Transforming a science classroom into a community of discourse requires the teachers to set a learning environment which stimulates and supports giving verbal explanation, comparing and critically evaluating different points of view on global warming issue. Some of the issues/dilemma for the group discussion can be, for example, how to cut global warming gas; laws for cutting emission of global warming gas; the possibility of cutting power plant and still having enough electricity etc. They, therefore, can talk about and make sound judgment regarding these socially embedded scientific issues. Through learning experience, higher order thinking and social responsibility is also cultivated in science classroom. Environmental responsiveness is the primary goal of environmental education. It can be enhanced by meaningful learning in which cognitive, motivational, social and contextual factors are activated.

Acknowledgement

This work was supported by the Thailand Research Fund (TRF) [grant number MRG5180147].

References

Abraham, M. R. & Wilkinson, V. M. (1994). A cross-age study of the understanding of five chemistry concepts. *Journal of Research in Science Teaching,* 31 (2), 147-165.

Boyes, E., & Stanisstreet, M. (1993). The Greenhouse Effect: Children's perceptions of causes, consequences and cures. *International Journal of Science Education, 15*(5), 531-552.

Corbett, J. & Durfee, J. (2004). Testing public (un)certainty of science: Media representations of global warming. *Science Communication*, 26(2), 129-151.

De Freitas, C.R. (2002). Are Observed Changes in the concentration of carbon dioxide in the atmosphere really dangerous? *Bulletin of Canadian Petroleum Geology, 50.*, 297-327

Dove, J. (1996). Student teacher understanding of the greenhouse effect, ozone layer depletion and acid rain. *Environmental Education Research. 2*, 89-100.

Driver, R., Squires, A., Rushworth, P. Wood-Robinson, V. (1994) *Making sense of secondary Science*. Routledge, London.

Francis, C., Boyes, E., Quslter, A. & Stanisstreet, M. (1993) Ideas of elementary students about reducing the greenhouse effect, *Science Education, 77*, 375-392.

Fisher, B. (1998). Australia student's appreciation of the greenhouse effect and the ozone hole. *Australian Science Teachers Journal, 44*(3), 46-55.

IPCC, (2001). Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change [Houghton, J.T., et al. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA,

Institute for the Promotion of Teaching Science and Technology (IPST), 2002. *National Science Content Standards*. MOE, IPST

Kilinç A, Stanisstreet M & Boyes E (2008) Turkish Students' ideas about Global Warming. *International Journal of Environmental and Science Education, 3*(2), 89-98

Mason, L., Santi, M. (1998), Discussing the greenhouse effect: Children's collaborative discourse reasoning and conceptual change.*Environmental Education Research*, Vol. *4*, 7-85.

McDonald-Gibson, C. (2007). Thai scientists fear global warming could empty world rice bowl. Retrieved on 9/12/2008 from http://www.terradaily.com/reports/ Thai_Scientists_Fear_Global_Warming_Could_Empty_World_Rice_Bowl_999.htmlPatanavanich, S.W. (1991) The basis for Thailand's response strategies to global warming. *TDRI Quarterly Review*. 6 (3), 7-12.

Taber, F. and Taylor, N. (2009) Climate of concern - a search for effective strategies for teaching children about global warming. *International Journal of Environmental & Science Education*,4, 97-116

Thomas Schelling, T. (2007). Developing countries will suffer most from global warming *Resource*. Winter,1-3

PART 2
ASSESSMENT OF STUDENT LEARNING AND DEVELOPMENT

ASSESSING TEACHERS' ORIENTATIONS TOWARD INQUIRY SCIENCE TEACHING: INSTRUMENT DEVELOPMENT AND INTERNATIONAL COLLABORATION

Renee' S. Schwartz, David Schuster, William Cobern & Brooks Applegate
Western Michigan University

Osman Titrek & Murat Iskender
Sakarya University

Abstract

This symposium describes the development of an instrument designed to assess teachers' orientations toward inquiry science teaching. Assessment items are generated by experienced teachers to provide case-based and problem-based situations. Items present a realistic teaching scenario for a science topic, pose a question about teaching strategy, and offer response options reflecting a spectrum of teaching orientations ranging from direct instruction through guided inquiry to discovery learning. Several rounds of item review and refinement involve expert panels and focus groups. Sets of validated items form a Pedagogy of Science Inquiry Teaching Test, with versions covering science topics relevant to US grades kindergarten through grade eight. A parallel instrument is being developed through collaboration with science educators and middle school teachers in Turkey. In addition to summative use, assessment items are useful formatively, with both future and practicing teachers, for problem-based learning of science inquiry pedagogy. The development model shows promise for wide-spread use through international collaborations.

Introduction

Science education standards in many countries reflect a growing commitment to the teaching of science as *inquiry* across the K-12 grades, i.e. instruction reflecting the investigative approach scientists use to discover and construct new knowledge. (AAAS 1990; NRC1996). Nevertheless there remains considerable educational and political debate about appropriate science instructional approaches, across a spectrum from 'direct' instruction through various degrees of guided inquiry to open pure 'discovery' learning. Standards documents advocate a guided inquiry approach that represents a range from more-to-less teacher guidance and less-to-more student responsibility (NRC, 2000).

Providing successful inquiry-based experiences in the science classroom demands a combination of science content knowledge and inquiry pedagogy knowledge, a combination included in a teacher's 'pedagogical content knowledge of inquiry science teaching.' During their preparation, pre-service teachers take science content courses and teaching methods courses. These teachers develop orientations toward science teaching that range from teacher-centered to student-centered, direct to open discovery. The goal of most teacher preparation programs is to develop teachers' orientations toward, and abilities of, inquiry science teaching, such that they are able to create, recognize and foster learning situations that develop scientific habits of mind and meaningful understanding of science and scientific inquiry. Preservice teachers' science content knowledge is regularly assessed, but it is just as important to assess their knowledge of and orientations toward inquiry pedagogy.

This paper presents the rationale and development of an instrument to assess science teachers' pedagogical orientations toward science inquiry teaching [POSITT]. This is a United States National Science Foundation funded project aiming to develop, validate, and pilot the instrument for use with preservice and inservice teachers of grades kindergarten through grade eight. An international component, also funded by the National Science Foundation, has enabled parallel development and validation of an instrument for use with Turkish science teachers of grades six through eight. The model of item development and enculturation has potential global applications.

This paper provides a summary of the project to day. We present (1) the need for teacher assessment and purpose of the POSITT; (2) the Science Teaching Orientations Spectrum model upon which items are based; (3) Item development; and (4) Item development and enculturation, exemplified by the collaboration between researchers and teachers in the US and Turkey. A more comprehensive paper with examples and more complete details of development and validation is available from the project director (david.schuster@wmich.edu).

Rationale

Pedagogical Content Knowledge and Teaching Science as Inquiry

To teach science successfully, teachers need to understand not only the science content but also how to translate the content and methods of science into appropriate instructional practices. Such ability is what Shulman called pedagogical content knowledge or PCK (Shulman 1986, 1987). PCK is the knowledge of effective instructional practices pertinent to specific content areas, and learner needs. For science teaching, PCK must emphatically include understanding of *inquiry* as an approach to the subject (Eick 2000; Lederman 1992, Lowery 2002). The POSITT is not an all encompassing measure of pedagogical content knowledge for inquiry teaching, yet it provides a critical piece of the picture. PCK for inquiry teaching must include a teacher's knowledge of best inquiry practice, in context, and reflect the teacher's orientations to identify inquiry instruction as "best practice."

Assessment of Content and Pedagogy during Teacher Preparation

During their undergraduate preparation, pre-service teachers commonly take several science content courses and a broad science teaching methods course. They are regularly assessed on their science *content* knowledge, but there is a critical need to complement this by assessing pedagogical knowledge of scientific inquiry. Currently this is rarely assessed because suitable instruments are not available. Preservice teachers do in fact have their teaching *practice* evaluated (Luft 1999), typically by classroom observation (e.g. Jenness & Barley 1999) and lesson plan review. While these measures are important to the development of teaching expertise, they are often in limited in scope. Observations and lesson plan review do not typically span the broad range of topics included across the science curriculum. Note that POSITT is *not* an observational instrument to evaluate inquiry teaching practice in the classroom. The POSITT aims to be a readily administered instrument for use during teacher preparation or in-service development programs.

Desired Assessment Instrument

An instrument to assess orientations toward science teaching should reflect the combination of science content knowledge and inquiry pedagogy knowledge that teachers require in practice. Thus, the instrument is a pedagogy assessment tool. Correspondingly, assessment items must be set in a context of relevant science topics and realistic teaching situations. For the desired purpose, the assessment items are both case-based and problem-based, requiring responders to *apply* their understanding of science pedagogy to determine appropriate teaching strategies in particular cases.

Various Summative and Formative Uses

The assessment can be used for several summative and formative purposes. Teacher educators need to evaluate how students moving through their preparation programs for science teaching are developing effective inquiry pedagogy knowledge that will form the basis for their practice. Researchers and educators will find such an instrument useful for evaluating and improving the effectiveness of science teacher education programs. While the instrument has potential for summative use, a real strength lies with its potential for formative use: (1) as an indicator of a teacher's tendencies toward inquiry pedagogy in the context of a variety of science topics, and (2) in providing classroom "cases" for discussion to promote understanding of inquiry pedagogy. Feedback from such assessment is known to be a very effective factor in learning (Black 2004). The role of observation of teaching practice in our project will be to field-validate the instrument's predictive validity.

Design Considerations for the Assessment

Science teacher education urgently needs an assessment instrument specific to the inquiry teaching of science, easily used, matched to grade range topics, validated by research, and able to serve summative and formative functions. This represents a new type of assessment, with no previous exemplars, so design characteristics needed to be thought out from scratch, often emerging simultaneously with initial item development, toward more systematic approaches and principles. The development process is explained in the next two sections.

Science teaching orientations spectrum model

An inquiry teaching orientations assessment tool should be based upon widely recognized views of appropriate inquiry teaching, and how this contrasts with other instructional approaches. There is a wide range of possible practices which might be called 'inquiry', depending on the degree of instructor guidance and student autonomy (NRC, 2000). For our purposes we adopt a 'guided inquiry' view of desired science instruction and make clear what we mean by specifying a model. This has the following features.

Scientific inquiry

The approach to a science topic should reflect characteristics of scientific inquiry; i.e. the processes and thinking that scientists use to explore the natural world and develop new knowledge. Students thus explore phenomena and collect observations toward formulating concepts and laws. This contrasts with simply being presented with the end-product knowledge commonly associated with direct instruction.

Guided

Students' learning paths should be appropriately structured and guided by the instructor. Guided inquiry can flow from more to less guidance (and back again), depending on the situation.

Science Teaching Orientation Spectrum

To fully understand what a guided inquiry approach is one must also know what it is not. Thus we need to contrast it with other possible science teaching approaches. The concept of 'science teaching orientations' is useful here. Teachers' orientations can range across a wide spectrum from direct instruction through guided inquiry to open discovery learning. For our purposes it is useful to consider four main orientations in this spectrum, labeled A, B C and D as follows.

↑ Science presented as a known product ('ready-made-science')

A. Direct didactic.

B. Direct interactive.

↓ Science as made through inquiry ('science-in-the-making')

C. Guided inquiry

D. Open (unguided) discovery

Note that two epistemologies are reflected here. There are two options within each, differing in degree. This spectrum view gives a natural basis for designing pedagogy assessment items, where response choices can reflect these orientations.

Assessment Item Design Considerations

Science Content and Level

The science context for a pedagogy assessment item is based on science topics specified in US national standards at elementary and middle school level (E.g. AAAS 1993: NRC 1996). Note that since we are testing science pedagogy (in context) rather than science content, many items may serve well for pedagogy assessment at high school level also.

Inquiry Instructional Stages

Our pedagogy assessment design is based on the model of guided inquiry instruction outlined above. Items thus need to have aspects that involve both scientific inquiry and a suitable degree of guidance, as reflected for example in a learning cycle approach.

Case-Based and Problem-Based Nature of Assessment

Teaching inquiry-based science successfully in the classroom requires a *combination* of science content knowledge and inquiry pedagogy knowledge. The knowledge also has to be *applied*, to deal with specific topics and cases. The assessment design has a number of features to reflect these needs. One feature is that items are posed not in generalities about inquiry but in terms of particular science topics in real teaching contexts. Thus items are case-based. They are also problem-based in that they present respondents with a practical problem to be solved, in the form of a *vignette* or *scenario* representing a realistic science teaching situation. There are different types of questions that can be posed, e.g. choosing a preferred teaching approach, evaluations of a teacher's actions so far, suggestions for what a teacher should do next, or ways of handling a question or classroom event. Suitable options are offered spanning a range of different approaches, based on the Science Teaching Orientations Spectrum. Such assessment reflects the kinds of instructional decisions that teachers have to make every day, both in lesson planning and on the spot in the classroom.

Basing a pedagogy assessment on concrete cases and problems has several advantages. Firstly, the assessment is more realistic and authentic. Secondly, it does not easily lapse into measurement of rote memory, nor of generalities about inquiry, nor guessing what might be "ideologically correct." Responding requires *application* and *evaluation*, in Bloom's Taxonomy terms (Anderson & Krathwohl, 2001). It requires both *understanding* of inquiry and *knowing how to use it* in new cases. Thirdly, because the assessment involves pedagogical approaches in real contexts, individual items are ideal for *formative* use in discussion and debate with pre-service students or in-service teachers. Our items are cast as 'problems' involving alternative pedagogical approaches to a given teaching situation. Working through such problems with future teachers operates as a scaffold for novices' current lack of schemas and serves as problem-based learning (e.g. Albanese, 1993) about inquiry in comparison to other instruction based on engagement with example cases.

Classes of Assessment Items

Soon after starting the task of creating this new kind of assessment, it became evident that two major classes of items were possible, which we call Class I and Class II. An item of Class I presents a case vignette and offers several alternative teaching approaches, asking the respondent to choose the preferred one. A Class II item describes one particular teaching approach, and asks the respondent to evaluate it. The structure of the two types is shown below. This item typology was crucial in guiding item creation and refinement.

Class I Item (Select preferred practice)	**Class II** Item (Evaluate a particular practice)
Given a particular teaching goal (for a specific topic) …	Given this particular teaching practice (for a specific topic) …
Which of the following teaching approaches would be best?	What is your evaluation of this teaching approach?

Formats of Assessment Items

For its intended wide-spread use the final instrument needs to be readily administered, easily scored and reliable, and hence objective item formats are desired. We devised two formats: the "Spectrum MCQs" (multiple choice) and "Likert Scale Testlets." It turns out that the former is well suited to items of Class I, while the latter is well suited to items of Class II. These two item formats are best explained through the specific example items given below.

1. Spectrum MCQ format

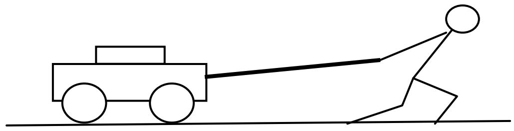
A Spectrum MCQ has four response options which correspond to four main teaching orientations in the spectrum described earlier. The example item below illustrates the complete structure with vignette, question and options.

Example 1

Teaching Force & Motion

A seventh grade teacher's aim is that students gain a qualitative understanding of the relationship between force and motion, namely that a constant applied force will cause an object to *change* its motion, i.e. accelerate (speed up or slow down, for straight-line motions).

A loaded wagon with little friction is useful in teaching this topic; one person can pull on the wagon with a constant force while others observe the motion.



Different teachers have different approaches to teaching the relation between force and motion. *Of the approaches below, which would be your first choice?*

A. Mr. Adams puts up a heading: Newton's second law of motion. Below it he writes 'a constant force causes acceleration' and explains the law carefully. He then demonstrates and verifies the law by pulling a loaded wagon in front of the class so that they see that it indeed speeds up.
B. Ms. Bell puts up a heading: Newton's second law of motion. Below it she writes 'a constant force causes acceleration' and explains the law carefully. She then has groups of students verify the law by pulling loaded wagons, following her instructions.
C. Ms. Campos raises the question: 'What kind of motion will result if there is constant force on an object?' She guides groups of students to explore this experimentally by pulling loaded wagons and observing what happens. From this evidence they then propose their own 'law' relating force and motion and write it on the board.
D. Mr. Doyle raises the general question of whether there is a relationship between force and motion. Groups of students are free to explore this in whatever way they wish, using any objects or equipment available. Groups are to report back at the end. Mr. Doyle does not prescribe what they should do, but helps with any equipment.

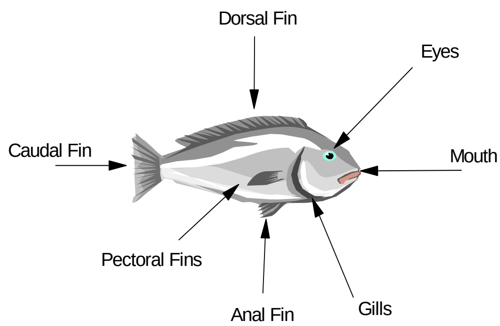
2. Likert Testlet Format

A Likert Testlet item has a very different format and different response structure. It also starts with a vignette, but then presents a set of statements it, each to be evaluated on a scale ranging from Strongly Disagree through to Strongly Agree. The example item below illustrates the format.

Example 2

Starting a lesson on fish

Mr. Lowe wants his fourth grade students to learn at a simple level about form and function, using a fish as an example. He begins the lesson by showing an overhead transparency of a fish. He names the various parts and labels them as shown.



Various comments about this way of starting the lesson are given below.

For each say whether you Strongly Disagree (SD), Disagree (D), Agree (A), or Strongly Agree (SA).

STATEMENT	SD	D	A	SA
1. Introducing the vocabulary the children will need for studying fish is a good way to start the lesson.				
2. It would be better to develop a sense of curiosity or questioning on the children's' part before presenting the information about fish.				
3. This is a good lesson so far because by learning the names of the fish parts, students are engaged and will ask questions about their function.				
4. It would have been better to start by letting students observe a real fish before showing a diagram of a fish.				

Note that a feature of the Likert scale format is that a set of different issues relating to a single vignette can be tested quite compactly; the set of issues comprises a Likert 'testlet'.

Because both formats are useful and have their own strengths, a complete POSITT instrument generally has two components, viz. a set of MCQ items and a set of Likert items. Having response options which reflect specified teaching orientations is an advantage for analysis and for identification of the teaching orientations of particular teachers, across a range of teaching scenarios. It can also show orientation variation and context dependence. From the set of items in the instrument one thus can build a profile of a subject's science teaching orientation(s), and see to what extent this can predict practice. Although objective-type tests are easy to administer and score, the creation and refinement of quality items with meaningful options is very demanding. Thus a substantial part of this project is the creative work of item development in a new assessment area.

Methods and Results

Item development and testing procedures

Item Writing Team

The US item writing team comprises four science education researchers, a doctoral research associate, and five experienced teachers representing grades kindergarten through eighth. Half of the group are also teacher educators. It is important to have teachers and teacher educators involved because they have direct knowledge of situations that teachers encounter and of the variations in their own and colleagues' ways of handling them. The teachers each generated up to 20 items based on their teaching experiences. The group met for 2-3 hour meetings every other month to discuss and refine items as well as discuss challenges and questions regarding item development. Item conceptualization and writing involved creating suitable classroom teaching scenarios, questions and response options. Items were usually based on personal teaching experience of topics, students and classroom events, as well as on knowledge of other teachers' practices and orientations. In the initial 'exploratory' item development stages, members of the writing team drafted 'concept' items, spanning a variety of science topics and the three grade divisions K-2, 3-5 and 6-8. They also aimed to cover the main stages of the science inquiry learning cycle, though it proved easier and more tempting to write for the initial 'set up' or exploratory stages. In this way we produced a first 'bank' of about 40 draft 'concept' items for consideration.

Face and content validity are established through the discussions. The research team is further refining items for use with focus groups.

Focus Groups and Review

Development includes a detailed feedback component for improving and validating individual items. A 'dissection' focus group of pre-service students and practicing teachers gives reasons for choosing or not choosing each option offered, as well as explaining their interpretation of the question. Project 2061 (AAAS 2007) uses a similar procedure for developing science *content* items, the detailed feedback proving essential to the item development and modification process. Groups of 3-6 preserive teachers and practicing teachers are assembled to review existing items. The focus group sessions further establish face, content, and construct validity of items. We utilize a two-step process with focus groups. First, the groups receive the vignette alone and are asked to write their own response to the probe. Comparison of free responses with existing choices support construct validity of the choices representing teachers' actual orientations. Choices that are not reflected in free responses are being refined. Secondly, groups are given complete items and asked to dissect the items and provide suggestions.

Thus far, we have held focus group sessions with about 20 preservice teachers enrolled in an elementary methods course. During the session, the students provided free-response answers to two of the items. They then responded and commented on each of the choices in the MCQ and Likert formats. A whole group discussion provided the most insights into the thinking of the preservice teachers as reflects their interpretation and rationale for selection of options. Results from the free-response data indicate the scenarios are realistic and reflect grade-appropriate teaching situations. The pre-written choices to class I items were all represented within the free-response data, providing validity to the options as representing teaching approaches the preservice population would find reasonable, while spanning the orientations spectrum. The discussion about the MCQ and Likert statements provided additional information on validity and reliability. Most respondents were consistent in their choices, either representing a more inquiry orientation or a more direct orientation. With few exceptions, we also found consistency with their open response statements. Rationale for selection indicated most statements were interpreted as intended. However, a few statements needed modification because of face validity issues. For example, one of the statements contained a question that the respondents found awkward and confusing. They avoided that option not because of the teaching approach represented, but because they did not like the questions the teacher posed. Only a few other statements were similarly problematic and refined accordingly.

We were able to identify inconsistent responses from a few participants. These inconsistencies provide additional information about the status of teaching orientations. For example, a few of the preservice teachers described very a teacher-centered direct strategy in their open response, yet chose a more guided inquiry strategy in the MCQ format. Such inconsistency suggests these preservice teachers can recognize an inquiry approach, but when asked to develop their own pedagogical ideas, they project a more teacher-centered orientation. In such cases, PCK for inquiry pedagogy is limited.

An 'expert' focus group, comprising people with recognized expertise in science teaching, science teacher education and/or assessment, was formed to review and critique both the criteria and the items developed. This includes an external panel of eight consultants from other institutions. Panel members rate items for science content, pedagogy and appropriateness of response choices. This feedback is used to establish the content validity of the items and their construct link to the criteria.

In-Class Piloting of Initial Items

Using a limited set of items, we have carried out some in-class pilots of the new assessment concept in an elementary science methods course in an institution with a large teacher education program. This try-out of individual items lends support to the viability of the concept. There was a concern that some students might sometimes guess the 'desired' answer to items, since guided inquiry is advocated in their preparation. Experience with these pilot items suggests that this will not be a significant problem. There was often a bimodal test distribution; methods students with good understanding of inquiry could readily identify it in problem-based teaching vignettes, while others were not easily able to identify guided inquiry approaches in contextual situations, nor able to guess it. At that stage, students seemed to either 'get it' or not, to a fair extent. This may be because at one end of the instructional spectrum, the idea that a teacher's role is 'telling' seems quite deeply entrenched; hence students with this orientation, conscious or unconscious, are likely to choose a corresponding response. At the other end, those with a very laissez–faire interpretation of inquiry will tend to choose 'discovery' responses. This pilot also fed into the improvement of items and options.

Final Instrument Composition

Our experience is that the task of creating vignettes, questions and response choices for pedagogy items is innovative, unfamiliar and far more challenging than producing items on science *content*. Thus besides producing the sets of assessment items we simultaneously devised typologies and guidelines for the new task. Our goal is to develop a total of about ninety objective pedagogy-of-inquiry items, thirty for each of the three grade groups, K-2, 3-5 and 6-8. Various POSITT instruments can be constructed from these; a test as administered to a particular group will have about 20 appropriately selected items comprised of both MCQ and Likert formats.

Validation of Final Instrument

The final part of the project will involve large-scale pilot testing, analysis, and validation of grade-appropriate versions of the complete compiled POSITT instrument. Testing and refinement involves two rounds of piloting with over a thousand students at several collaborating institutions. Besides construct-related validity, we are also interested in predictive validity, and will conduct subsequent blinded classroom observation studies to investigate the extent to which POSITT performance is a predictor of teaching practice.

Item development and enculturation: A collaboration between the United States and Turkey

In October 2008, members of the research team in the US met with two researchers at a university in Turkey and five Turkish middle school teachers (four science teachers, 1 English teacher). The purpose of the meeting was to begin a collaborative project to develop POSITT items for use with Turkish science teachers of grades 6-8. The international collaboration has generated multiple items for use with Turkish as well as US teachers. Our model for item development and cross-cultural sharing is underway, and shows promise for application within other cultures and languages. The process of item development involves researchers and teachers from Turkey and the US.

Item development for Turkish POSITT

This international collaboration for the project started in the US, when two Turkish academicians were visiting scholars at WMU. Before they returned to Turkey, the research project team discussed and compared Turkish and US education systems and science education programs. Turkish academicians were informed about current US Project of POSITT by American research team. Upon returning to Turkey, a POSITT research team was assembled in Sakarya Univeresity. The Turkish research team consists of four middle school science teachers (grades 6-8) and an English teacher. Several US items were translated by the Turkish -English teacher who would work with the research team on translation and validation of the Turkish instrument. The team reviewed all information about this Project in preparation for a visit from the US team to the Turkish site.

In October 2008, two members of the US research team visited Sakarya University and met with the Turkish researchers and teachers. This meeting involved visiting and observing two middle schools in Turkey, and one of the US researchers taught a science lesson to Turkish 6th grade students. These experiences in Turkish schools and with Turkish children provided a useful context and understanding for the development of items for the POSITT Turkish version. US and Turkish research teams and Turkish teachers met to examine the framework, Orientations Spectrum, and sample US items. US items were translated by a Turkish English teacher and reviewed by four Turkish middle school science teachers (grades 6-8). Discussions of the items centered on clarity of language, appropriateness of science content for Turkish system, and appropriate use of examples.

Enculturation of items

The inclusion of the teachers in item development and review is a critical feature of the model. Another critical feature was the visitation and observation of Turkish science classes. To develop inquiry-based scenarios that reflect the appropriate context, the writing teams certainly must come to understand the perspectives, environments, learners, and goals. The US items are developed by American science teachers, considering what is appropriate and typical of American students, classrooms, and science curriculum. Once translated into Turkish, the Turkish teachers aid revisions of items to reflect Turkish students, classrooms, and science curriculum. In essence, they facilitate the "enculturation" of items for use with Turkish science teachers. Items generated by the Turkish team will likewise be reviewed and select items will be enculturated for use with American teachers.

The example below presents both the US version and the Turkish version of an item developed in collaboration between the US team and Turkish team.

Turkey generated items

To date, the Turkey research and teaching team has generated 20 items for the middle school level Turkish POSITT. They are translated into English for review by the US researchers and teachers. These items will also undergo focus group review with Turkish teachers (Turkish version) and American teachers (English version). The focus group sessions will reflect the two-stage process previously described. The intention is to collect several items that can be used on both the English version and Turkish versions.

US generated items

The US research and teaching team continue to develop items. Select items relevant for Turkey grades 6-8 are translated and sent to the Turkish team for review and enculturation. Focus groups aid in the process of review and revision. Revised items may then be back-translated to English for review by the US team and focus groups.

Pilot testing

Selected items for use with Turkish teachers will be piloted in teacher preparation courses at the [Turkish] University. Finished items will be used for summative and formative purposes, as paralleled by the POSITT, US version.

Teaching Orientation Spectrum	ITEMS AS ENGLISH (Maddelerin İngilizce'si)	ITEMS AS TURKISH (Maddelerin Türkçe'si)
Scenario **(Senaryo)**	The aim of the seventh grade teacher Mrs. Aysegul is to make students gain qualitative knowledge about states of matter	Öğretmen Ayşegül'ün amacı öğrencilere maddenin halleri ile ilgili nitelikli bilgi kazandırmaktır.
DIRECT DIDACTIC (Doğrudan öğretim)	Teacher put up a description about the states of matter gives some examples such as water, sponge, stone and balloon blown up and asks students to put them in these 3 (three) categories	Öğretmen maddenin hallerinin -katı, sıvı, gaz- tanımı yapar ve su, sünger, taş ve şişirilmiş balon örnek olarak verir ve bunları öğrencilerin bu 3 (üç) kategoriye koymalarını ister.
DIRECT ACTIVE (Doğrudan etkileşim)	Teacher gives description of the states of matter' solid, liquid and gas. The teacher passes out students the samples: water, sponge, stone and balloon blown up and ask them classify according to the states of matter.	Öğretmen maddenin hallerinin -katı, sıvı, gaz- tanımı yapar. Öğretmen öğrencilere su, sünger, tas, şişirilmiş balon örneklerini öğrencinin önüne koyar ve sınıflamasını ister.
GUIDED INQUIRY (Yönlendirilmiş araştırma)	Teacher passes out samples of water, sponge, stone and balloon to the students then raises the question: What are some differences in how these objects look and feel? Students explore objects then show their ideas. The teacher uses these ideas to classify states of matter.	Öğretmen öğrencilere su, sünger, tas, şişirilmiş balon örneklerini öğrencinin önüne koyar ve soruları çoğaltır: Bu nesnelerin nasıl göründüğü ve hissedildiğine ilişkin farklılıklar nelerdir? Öğrenciler nesneleri keşfeder ve fikirlerini belirtir. Öğretmenler bu fikirleri maddenin hallerinin sınıflamasında kullanır.
OPEN DISCOVERY (Açık buluş)	The teacher sets groups of students. Each group is provided the samples of water, sponge, stone and balloon. One of the members of each group has blind fold. He \ she tries to feel the samples and the each group comes up with ideas about the similarities and differences of the samples. Then as a whole class they share the ideas and come to the conclusion about the states of matter.	Öğretmen öğrencileri gruplandırır. Her gruba su, sünger, taş ve balon örnekleri sağlanır. Her gruptan birinin gözleri bağlanır. O maddeleri hissetmeyi dener ve her grup örneklerin benzerlik ve farklılıklarını anlamaya çalışır. Sonra bütün olarak sınıf, fikirleri paylaşır ve maddenin halleri hakkında sonuca varırlar.

Conclusions and Implications

At the conclusion of the project, a case-based objective instrument for assessing pedagogical orientations toward inquiry science teaching will have been developed and validated for use in American teacher education. A similar instrument will be useful for science teacher education in Turkey. Once complete, POSITT versions and the development model will be made widely available. Thus teacher educators at any undergraduate institution will be able to assess the effectiveness of their instruction and development of future teachers' orientations toward inquiry teaching, and be in an informed position to improve their science teacher education programs. Although the US POSITT is primarily designed for pre-service K-8 teachers, items of this nature have the potential for much broader uses, e.g. in-service, or at secondary level, or for graduate students and teaching assistants in science education programs. We are designing the POSITT Turkish version to reflect the needs of middle school teacher education. The POSITT development model has wide-spread potential. We welcome international partners for continued collaboration.

Acknowledgement

The project is supported by National Science Foundation grant DUE-0512596. Items and detailed description of framework, item development and use can be obtained from the Project Director, David Schuster (david.schuster@wmich.edu).

References

Albanese, M., & Mitchell, S. (1993). Problem based learning: A review of literature on its outcomes and implementation issues. Acad Med, 68(1), 52-81.

American Association for the Advancement of Science (AAAS). (1990). Project 2061: Science for all Americans. Washington, DC: American Association for the Advancement of Science, Inc.

American Association for the Advancement of Science (AAAS). (2007). Getting assessment right – rigorous criteria and student data inform Project 2061's assessment design. 2061 Today, Winter 2007, Vol 17, No 1, pp 1, 4, 5.

Anderson, L. W., & Krathwohl, D. R. (2001). A taxonomy for learning, teaching, and assessing: a revision of Bloom's Taxonomy of educational objectives. New York: Longman.

Black, Harrison, Lee, Marshall & Wiliam 2004, Assessment for Learning - Putting it into practice, Open University Press.

Eick, C. J. (2000). Inquiry, nature of science, and evolution: The need for a more complex pedagogical content knowledge in science teaching. Electronic Journal of Science Education, 4 (3).

Karplus, R. (1977). Science teaching and the development of reasoning. Journal of Research in Science Teaching, 14(2), 169-175.

Lederman, N. G. (1992). Do subject matter knowledge, pedagogical knowledge, and pedagogical content knowledge constitute the ideal gas law of science teaching? Journal of Science Teacher Education, 3(1), 16-20.

Lowery, L. F. (2002). The nature of inquiry. Science Link, 13(1), 3.

Luft, J. A. (1999). Assessing science teachers as they implement inquiry lessons: the extended inquiry observational rubric. Science Educator, 8(1), 9-18.

National Research Council. (1996). National science education standards. Washington, DC: National Academy Press.

National Research Council. (2000). Inquiry and the National Science Education Standards: A guide for teaching and learning. Washington, DC: National Academy Press.

Shulman, L. S. (1986). Those who understand: A conception of teacher knowledge. American Educator, 10(1), 9-15.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. Harvard Educational Review, 57(1), 1-22.

Tunç, T., Agalday, M., Akçam, H. K. ve diğ.(2008). *İlköğretim fen ve teknoloji 6. sınıf ders kitabı.* (Science and technology book for 6 th grade primary schools). Ankara.

Tunç, T., Bağcı, N., Yörük, N., ve diğ.(2007). İlköğretim fen ve teknoloji 7. sınıf ders kitabı (Science and technology book for 7 th grade primary schools). Ankara.

Tunç, T., Bakar, E., Başdağ, G., ve diğ.(2008). İlköğretim fen ve teknoloji 8. sınıf ders kitabı (Science and technology book for 8 th grade primary schools). Ankara.

SELF-DIRECTION, MOTIVATION AND CONCEPTUAL CHANGE IN PHYSICS LESSONS IN SECONDARY SCHOOLS

Angelika Wolf & Matthias Laukenmann
Universitiy of Education, Heidelberg

Abstract

In order to meet the demands of a life-long scientific learning, the curriculum for secondary schools in Germany calls for the schools to set a solid foundation for science education within its schooling. The meaning of motivational orientations for the learning in the mathematical-scientific lessons (Renkl 2002), the learning-supportive teaching procedures and teaching methods are in the centre of this research project according to Hentig's postulate, that learning requires independence individual responsibility and self-control of each learner H.v. Hentig 2004), the conditions for an adequate motivation of learning and, by means of that, for successful and lasting learning can be created by the experience of competence, autonomy and relatedness. In their Self-determination theory of motivation (SDT) Deci and Ryan (1993) present a possible theoretical frame. This project tries to investigate the sustainability of this postulate for physics lessons at secondary schools. We are testing from three perspectives (students, teachers and raters), whether or not an observable connection between the degree of autonomy given (by the teachers) within lessons and the students' subjective experience of social integration, autonomy and competence, exists. Therefore the study aims at investigating, whether the realization of those basic needs possibly lead to different levels of motivation and whether or not the degrees of self-determination, motivation and cognition are correlated using multivariate analysis methods.

Introduction

Based on the postulate of Hentig and the self-determination theory of motivation by Deci and Ryan, we come to the following aims of research, see figure 1.

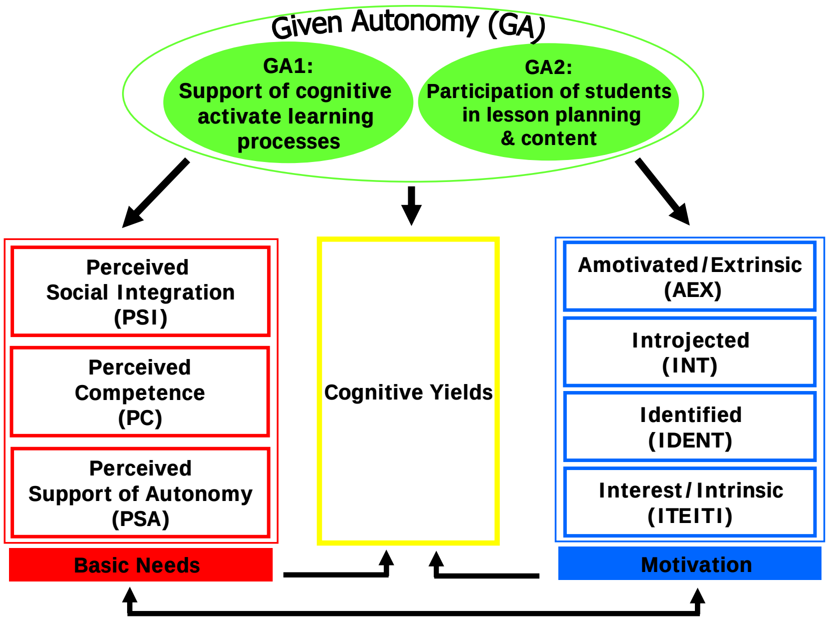


Figure 1. Aims of research

We like to investigate all the correlations, which in this diagram are presented as darts. The first results of the pilot study, which took place from March to June of this year are presented in this article.

Rationale

Deci and Ryan (1993) suggest that learning motivation being based on self-determination has positive effects on the quality of learning. In connection with the above mentioned postulate by Hentig this means, that the fulfillment of basic needs also leads to a greater success of any learning process. According to Deci and Ryan (1993) basic needs are social relatedness, autonomy and competence, which every subject aspires by nature. Consequently, lower success in learning has to be expected when these basic needs remain unfulfilled.

Apart from other motivational theories, SDT is especially suited for the task at hand, because it offers a refined description of extrinsic motivational orientation, which presents itself as a good foundation for an in-depth analysis of the learning behaviour of pupils, and allows for an exact allocation of the pupil's motivation and learning behaviour. This is of great value in everyday life at school, because processes of teaching and learning themselves are highly determined by extrinsic requirements[1] towards the learners. The determination of success of learning respectively the growth in learning is defined in the sense of the conceptual change theory, based on a moderate constructivist oriented theory of cognition (Gerstenmaier, Mandl 1995; Labudde 2000).

Grounded in this theoretical framework the following research questions arise:

1.) Do the learners experience themselves according to the lesson's degree of autonomy, individual responsibility and self-determination to a different extent autonomous, competent and socially integrated in terms of the SDT?
2.) How does the amount of autonomy, individual responsibility and self direction correlate with the learning success regarding the development of scientific concepts in lessons?
3.) Are there correlations between the „Given Autonomy" within the lesson and the different aspects of motivation?
4.) What connection can be observed between the satisfaction of basic needs and the different qualities of motivation on the learner's side within the videographed lessons?
5.) Are significant correlations between the experience of autonomy, motivational orientation and cognitive yields provable?

Methods

The study has two phases, which share a common design.

The Pilot Study took place from March to June of this Year. The start of the main study takes place in October 2009. The sample for the studies contains altogether about 30-40 teachers and 600-1200 pupils of the eighth grade at secondary schools' science lessons in Germany.

The data collection both in the pilot and the main study investigates "instruction form" with a focus on the degree of autonomy, individual responsibility and self-determination offered to the learners. The data is collected multiperspectively:

1.) Assessment of the lesson criteria by pupils: Use of standardized and validated charts from COACTIV (Jordan et al 2006), TIMSS (Baumert et al. 2000) and BIJU (Baumert et al. 2001).

[1] The school's tasks are mainly determined by the external factor oft he national curriculum.

2.) Direct teaching observation of several teachers of the secondary school in physics lessons focussing on their instruction forms, as well as video recording of the teaching events. The evaluation of the instruction forms, as well as the evaluation of the video recordings occurs on the basis of standardised scales (COACTIV; DIPF (Klieme et al. in 2005); PISA 2000/2003 (Baumert et al. in 2003), video-studies of the IPN (Seidel et al. in 2003)).

3.) Questioning teachers towards their perceptions about the self-assessment independence in the lessons and to the teaching quality with the help of a standardised questionnaire (also COACTIV (Jordan et al in 2006); DIPF (Klieme et al. in 2005); PISA in 2000/2003 (Baumert et al. in 2003)).

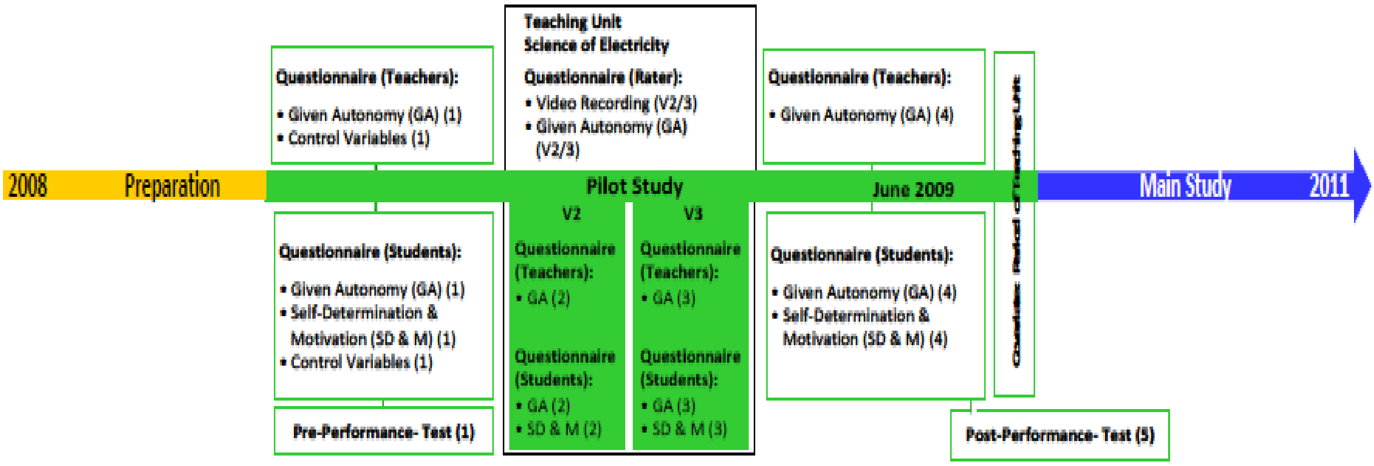


Figure 2. Design of the Pilot Study

Moreover, at the beginning controlling variables are raised, as for example professional interest and material interest of the pupils, perceived technical, didactic and educational competence of the informing teacher, reported mark in physics in the preceding school year or well-chosen areas of the teacher's personality which have an influence on the learning success.

The factors experience of relatedness, of competence, of autonomy and motivational orientation in the sense of the self-determination theory are raised with the help of a standardised questionnaire of the interest study IPN (Seidel et al. 2003, Prenzel et al. 1996; Rimmele et al. 2001) at the beginning as well as at the end of the teaching unit.

To measure the cognitive learning success (previous and specialized knowledge) two knowledge tests at the beginning and at the end of the series of lessons are conducted.

The evaluation of the data encompasses different analysis procedures. In order to be able to answer research questions 1-4, which investigate the degree of self-determination, success in learning and motivation in the lessons dependent on the participating teacher were subject to our research focus, a quantitative analysis based on multivariate methods (covariance analysis) will be executed.

In research question 5 we try to find out whether between the experiences of autonomy, the motivational orientations and the cognitive yields empirically significant correlations can be observed. In order to do so also a correlation analysis with multivariate methods will be applied.

Besides a path analysis is aimed to the check of the accepted cause-effect respect between the instruction form and the specialized knowledge of the pupils. Moreover, a multi-level effect model will be developed in order to estimate the upraised data at an individual level as well as at class level

Results

We called for eight teachers and their classes (about 250 students) mainly to test our research design with the questionnaires and the video analysis. For the data acquisition, we visited one specific class several times .

We used standardized scales and one self-developed scale to acquire the different aspects. The first application aimed at getting the control variables of teachers and of the students and also their opinion about given autonomy within lessons. The students were also asked about their basic needs and motivation. The Pre- and Post-Performance-Tests weren't conducted, because the teachers had different topics in their teaching units. In the main study, which starts in October this year, all teachers are teaching electricity.

During the teaching units, at the beginning and at the end, we visit classes two times to film the lessons and to rate them with the same questionnaire about given autonomy that the students and teachers get to fill out. By these means we obtain three perspectives. The students were also asked with a short test for their basics needs and motivation.

At the end of the teaching unit, we asked the students and teachers for Given Autonomy and Self – Determination and Motivation again.

In the main study we also acquired as a covariate the period of teaching unit for comparison.

Confirmation of the questionnaire-reliabilities

With the data, we could on the one hand confirm the reliabilities of the standardized scales used (α = .68 to .94). On the other hand our self-developed scales showed a Cronbach's α from .74 to .80.

The scales for acquisition of the "Given Autonomy" (GA) were self-developed and tested in a heuristically and defined way.

With regard to content and empirical, the optimum for the scale GA was a two-way solution with GA 1 „Support of cognitive active learning processes" and GA 2 „Participation of students in lesson planning and content".

Multi-perspective design

To collect the construct „Given Autonomy in lessons" (GA 1/2) in a comprehensive way, we chose three perspectives (teachers, students and raters). To compare the results, we constituted the same questionnaire but parallelized with regard to content.

In Figure III, you can see the Given Autonomy estimated differently by Teachers, Students and Raters within the lessons. At GA1, the teachers, expect the teacher 1, up rate as the students and rater. Maybe, the fact "Social-Desirability-Response-Set" plays a decisive role. At GA 2, there no such differences are to be noticed.

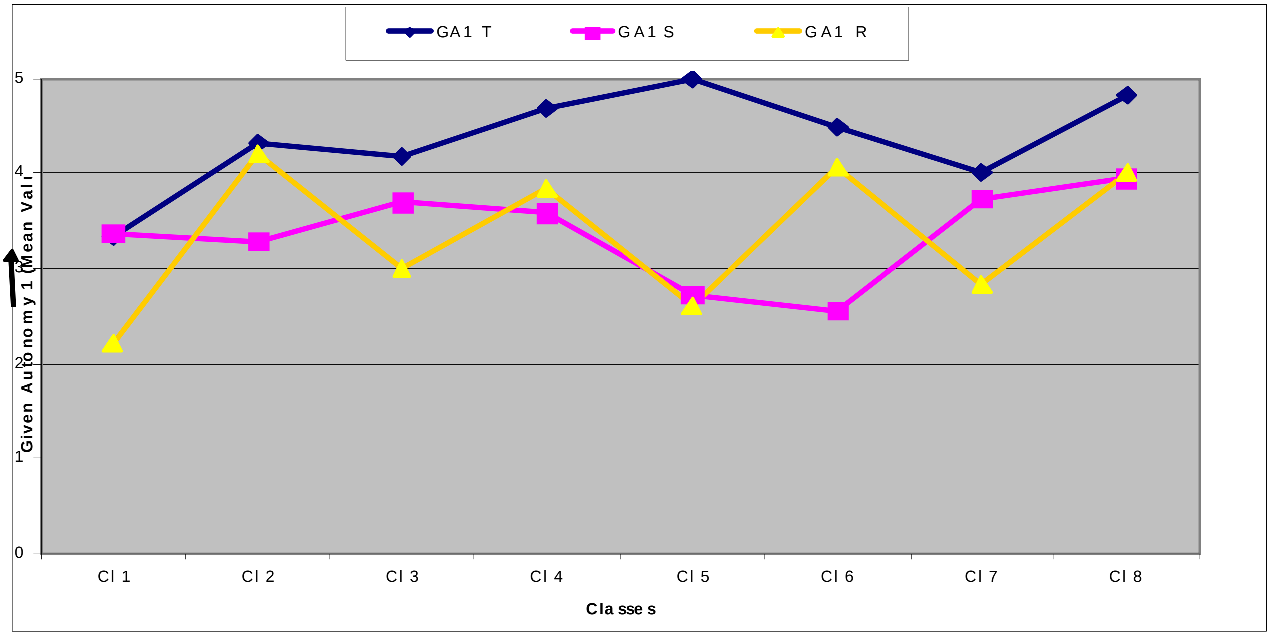


Figure 3: Multi-perspective design (GA1)

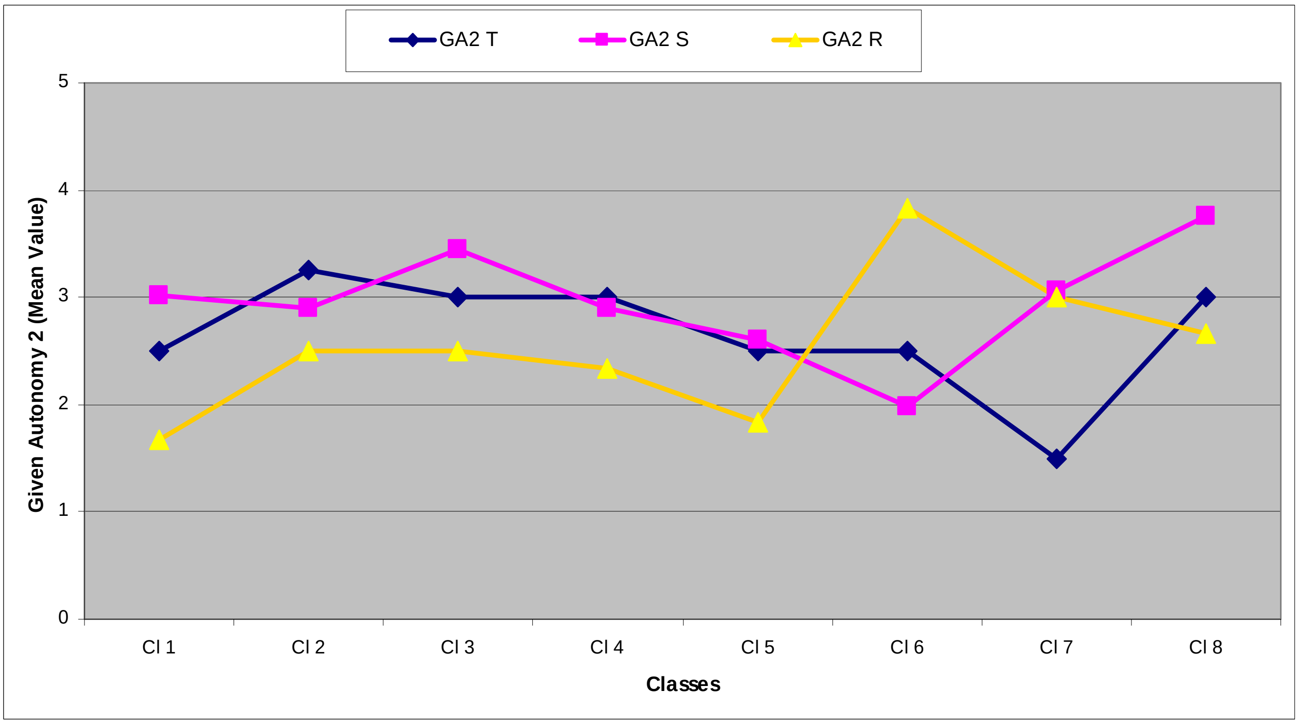


Figure 4: Multi-perspective design (GA2)

Answer to the Research Question 1

Do the learners experience themselves according to the lesson's degree of autonomy, individual responsibility and self-determination to a different extent autonomous, competent and social integrated in terms of the Self-Determination Theory of Motivation (SDT)?

Correlations				
		PSI	PC	PSA
GA1 Support of cognitive activate learning processes	Correlation according to Pearson	,38**	,37**	,51**
	Significance (bilateral)	,000	,000	,000
	N	194	194	194
GA 2 Participation of students in lesson planning and content	Correlation according to Pearson	,32**	,32**	,42**
	Significance (bilateral)	,000	,000	,000
	N	194	194	194

Figure 5: Correlation GA1/2 and PSI/PC/PSA

For the first answer to this research question, the two subscales GA 1 and GA 2 were correlated with the Perceived Social Integration (PSI), the Perceived Support of Autonomy (PSA) and the Perceived Competence (PC).

Between scales GA1 and GA 2 and the basic needs (PSI, PSA, PC) the data show high significances and strong correlations. The correlations are according to the SDT – there are no negative correlations. Remarkable is the strong correlation between GA 1 and GA 2 and the Perceived Competence.

Answer to the Research Question 3

Are the correlations between the "Given Autonomy" within lessons and the different aspects of motivation?
First, we calculate the mean values of the standardized scales to acquisition the motivation Amotived/Extrinisic (AEX), Introjected (INT), Identified (IDENT), Interest/Intrinsic (ITEITI)) and graph them. Afterwards the values of the classes were compared with the mean value of all participating classes.

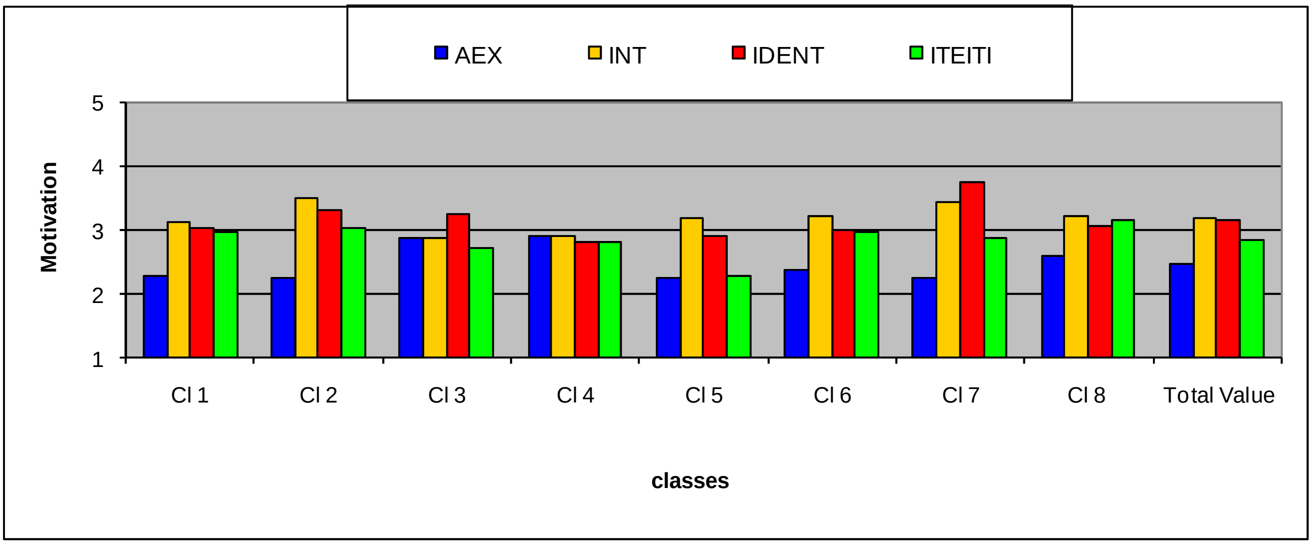


Figure 6: Comparison of motivation in several classes

The concrete correlation between the "Given Autonomy" and the different aspects of motivation is shown in the results in figure 7.

Correlations					
		AEX	INT	IDENT	ITEITI
GA1 Support of cognitive activate learning processes	Correlation according to Pearson	-,09	,13	,22**	,28**
	Significance (bilateral)	,174	,052	,001	,000
	N	228	228	228	228
GA 2 Participation of students in lesson planning and content	Correlation according to Pearson	,04	,11	,13*	,18**
	Significance (bilateral)	,546	,092	,044	,006
	N	228	228	228	228

Figure 7: Correlations GA 1/2 and AEX/INT/IDENT/ITEITI

There is a high significance between the scales GA 1 and the identified and intrinsic motivation. Between GA 2 and identified motivation, there is significance and a high significance between GA 2 and the intrinsic motivation.

Maybe the results of the main study will give rise to more detailed information.

Answer to the Research Question 4

What connection can be observed between the satisfaction of basic needs and the different qualities of motivation on the learner's side within the videographed lessons?

Correlations				
		PSI	PC	PSA
AEX	Correlation according to Pearson	,38**	,37**	,51**
	Significance (bilateral)	,000	,000	,000
	N	194	194	194
INT	Correlation according to Pearson	,32**	,32**	,42**
	Significance (bilateral)	,000	,000	,000
	N	194	194	194
IDENT	Correlation according to Pearson	,35**	,43**	,45**
	Significance (bilateral)	,000	,000	,000
	N	194	194	194
ITEITI	Correlation according to Pearson	,40**	,55**	,44**
	Significance (bilateral)	,000	,000	,000
	N	194	194	194

Figure 8: Correlations between Motivation and the Basic Needs

The results of the correlations between the different aspects of motivation and the basic needs are highly significant and thus also seem to confirm the theory.

A student, who doesn't perceive him- /herself as socially related (PSI), is with high probability amotivated and extrinsic motivated. (AEX). The other way around, a student, who does perceive him- /herself as social related, is in the trend introjected (INT), identified (IDENT) or intrinsically motivated (ITEITI). It is important to understand, that according to Deci and Ryan, all the three basic needs have to be satisfied in order for a student to have a high quality of motivation.

References

Deci, E. L. & Ryan, R. M. (1993). Die Selbstbestimmungstheorie der Motivation und ihre Bedeutung für die Pädagogik. Zeitschrift für Pädagogik 39, 223-238.

Deci, E. L. & Ryan, R. M. (2000). The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behaviour. Psychological Inquiry, Vol. 11, No.4, 227-268

Duit, R. (2000). Konzeptwechsel und Lernen in den Naturwissenschaften. In Duit, R. & v. Rhöneck, Ch. (Hrsg.), Ergebnisse fachdidaktischer und psychologischer Lehr-Lern-Forschung. Kiel: IPN.

Gerstenmaier, J. & Mandl, H. (1995). Wissenserwerb unter konstruktivistischer Perspektive. Zeitschrift für Pädagogik 42, 867-888.

Helmke, A. (2003). Unterrichtsqualität erfassen, bewerten, verbessern. Seelze: Klett

Labudde, P. (2000). Konstruktivismus im Physikunterricht der Sekundarstufe II. Bern: Haupt.

Ministerium für Kultus, Jugend und Sport Baden-Württemberg (Hrsg., 2004). Bildungsplan 2004. (www.bildung-staerkt-menschen.de)

Pintrich, P. R., Marx, R. W. & Boyle, R. A. (1993). Beyond cold conceptual change: the role of motivational beliefs and classroom contextual factors in the process of conceptual change. Review of Educational Research 63, 167-199.

Posner, G., Strike, K., Hewson, P. & Gertzog, W. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. Science Education, 66, 211-227.

Renkl, A. & Stern, E. (1994). Die Bedeutung von kognitiven Eingangsvoraussetzungen und Lernaufgaben für das Lösen von einfachen und komplexen Textaufgaben. Zeitschrift für Pädagogische Psychologie 8, 27-39.

AN ENERGY-BASED MODEL FOR TEACHING THE CONCEPT OF ELECTROMOTIVE FORCE. STUDENTS' DIFFICULTIES AND GUIDE LINES FOR A TEACHING SEQUENCE

Jenaro Guisasola
Univeristy of the Basque Country

Antonio Montero
Secondary School Institute. Nerja

Abstract

Our work deals with the difficulties which arise at university level in the analysis of the operation of simple direct current circuits which includes the concepts of energy and electrical current. The final aim of this study is to provide a set of guidelines for designing teaching sequences to serve as guidance in teaching the interpretation of energy and proof of the movement of charges between two points on a circuit (concept of potential difference) and throughout a circuit (concept of electromotive force).

In order to investigate students' difficulties in understanding, a questionnaire based on an analysis of the theoretical and epistemological framework of physics was used. It was put to 3rd course physics students undergoing examinations. The results of the study show that students' difficulties seem strongly linked to the absence of an analysis of the work carried out on the circuit and its energetic balance. In this regard, most 3rd course physics students still do not clearly understand the usefulness of concepts of potential difference and fem. This leads to the necessity to design tasks and problems which provide students opportunities to understand an energetic model which involves the concepts of potential difference and electromotive force concepts.

Introduction

Our work deals with the difficulties which arise at university level in the analysis of the operation of simple direct current circuits which includes the concepts of energy and electrical current. Mulhall et al. (2001) conclude that teachers have problems describing energy transformations in a circuit within a coherent framework and admit that they do not have a qualitative idea of the concept of potential difference and tend to avoid it in their explanations. The final aim of this study is to provide a set of guidelines for designing teaching sequences to serve as guidance in teaching the interpretation of energy and proof of the movement of charges between two points on a circuit (concept of potential difference) and throughout a circuit (concept of electromotive force).

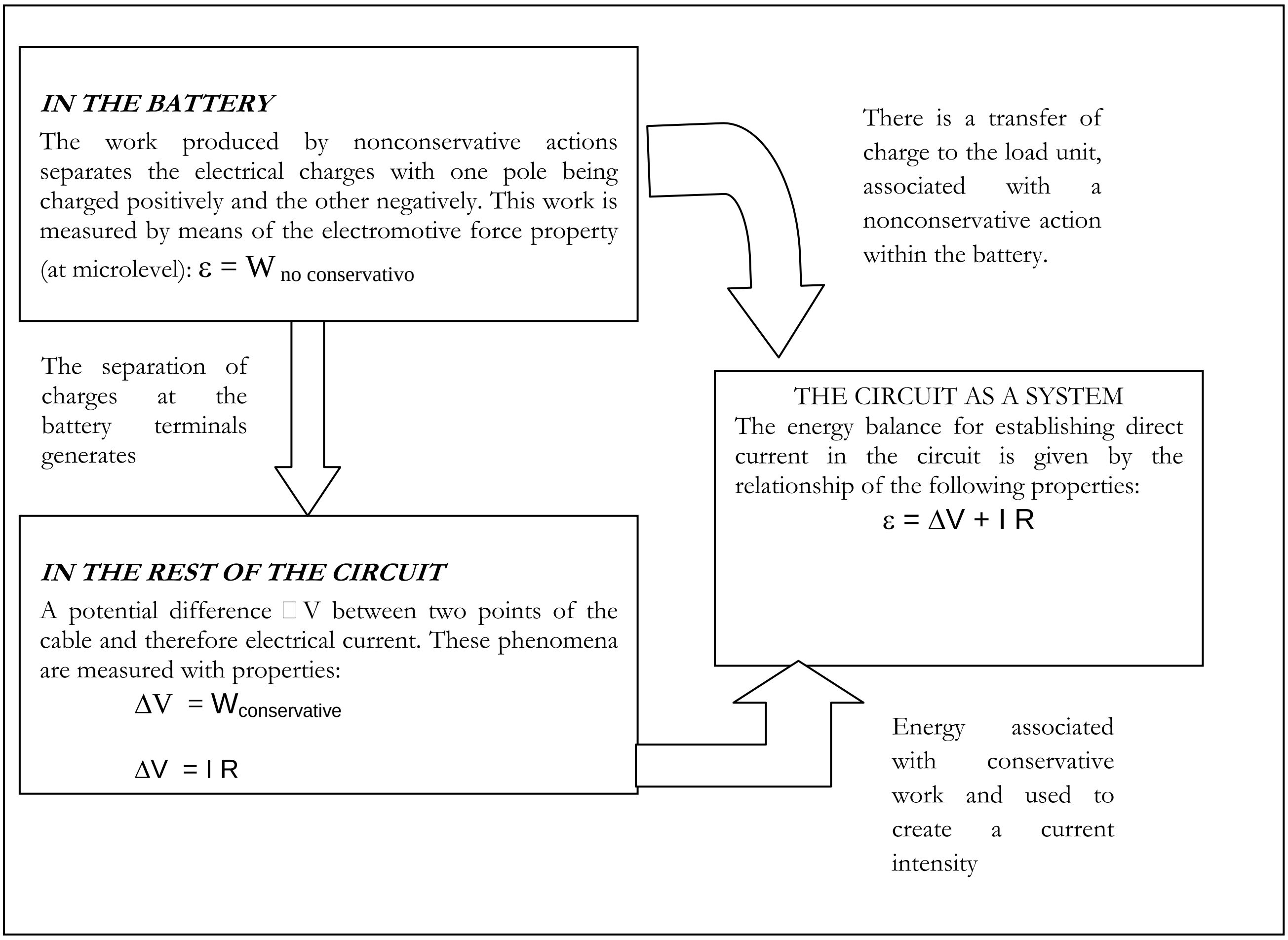
A set of interrelated reasons has converged in the choice of the topic of electromotive force (emf) in the context of direct current circuits, about the teaching-learning of which there is little research (Guisasola et al. 2005). Firstly, this notion is included in Secondary School programmes (age 16-18) and first-year engineering and science university courses. Secondly, it is a basic prerequisite for explaining the functioning of a direct current circuit. Thirdly, their principal technological innovation is the battery, the subject being so spectacular that one can now scarcely conceive a society without them; pacemakers, hearing aids, mobile telephony, a great number of home appliances...etc.

Rationale. The decisive role of “work done in the circuit”

We took the notion of ‘learning indicator’ as a reference for our content analysis. The concept of 'learning indicators' allows us to sequence the principle stages that teachers must work out when designing teaching programmes (Guisasola et al 2008). We have also used this concept in order to analyse the most significant ideas and forms of reasoning that make up students' understanding.

From the physics point of view, it is well known that a potential difference must exist between two points of a conducting wire, in order for the charges to be displaced along the wire. One way to generate a potential difference is to separate charges of different polarity within a spatial area, and in the case of a direct current circuit this function is realised by the battery. In this context of simple electrical circuits, electromotive force is a property that quantifies the energy delivered to the charge unit by the electrical generator. A series of 'non-electrostatic electrical actions' takes place in the battery, through which energy is delivered to the charge unit and this energy is quantified by means of the property 'electromotive force'. Therefore, it is the “work done” to produce and maintain the electrical current which determines its relevance to the analysis of the movement of charges on a simple continuous current circuit. Whereas the potential difference measures the work used up by the charges when moving from one point to another on a circuit (work carried out by conservative forces), the fem measures the work carried out by the generator to generate a potential difference by separating charges (work carried out by non-conservative forces).

These considerations configure a universe of concepts where the idea of electromotive force in relation to other concepts that conform the explanatory model of how and why continuous electric current is generated in a simple circuit, acquires significance. Table 1 presents these relationships:



From the definition of the explicative model a new question arises, Are this concepts resistant learning difficulties to the usual teaching? To answer this question we carried out an empirical research which we explain in the next section

Methods

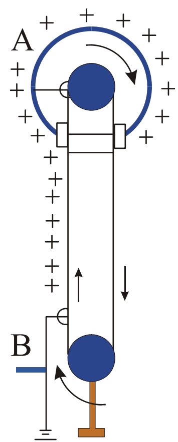
In order to investigate students' difficulties in understanding, a questionnaire based on an analysis of the theoretical and epistemological framework of physics was used. It was put to 3rd course physics students undergoing examinations.

The analysis of the answers was based on the same criteria which we used to make the analysis of the physics contents referred to in the preceding section. The students' answers were examined independently by two members of the research group who sought similarities and differences and chose significant statements, comparing them in order to obtain cases of agreement or variations, and then grouping these statements into categories (Watts et al. 1997). After the initial categories had been established the researchers met to discuss those which each had found and to check them until a consensus on final categories was reached. The questionnaires were then reviewed again to determine whether the types of answers adequately described and indicated the data.

Results

The questionnaire was completed by 56 3rd course physics students at the University of Granada. These students had passed physics in the first course and electromagnetism in the second, in addition to other subjects connected with the technological applications of electromagnetism.

Question Q2, the aim of which was to investigate student' understanding of the concepts of potential difference and emf within a context not explicitly connected with the normal teaching of electrical circuits, is given by way of example.



Q2. As you know, in the Van der Graaf generator rubbing a belt of insulating material in movement gives rise to a separation of charges (negative in the metal sphere (zone A) and positive in zone B (see figure). If a conducting cable is placed between zone A and zone B, two students made the following prediction:

S1: An electrical current will be produced due to the potential difference between both zones.

S2: There is no electromotive force in the generator as it is not a battery and therefore there will be no current.

Which of the above do you think is the correct prediction? Justify your answer.

46% (N= 26) of the answers said that student S1 had made the correct prediction, but only 16% (N= 9) explained that the cause of the electrical current was the potential difference between two points. Most of these answers (N=17) explained that there would be current due to the differences in "amounts of charge" accumulated in both zones (N=12), or because charges of a different kind accumulated and would be attracted (N=5). Of the 9 answers which correctly justified student S1's prediction, five (9%) also explained that student S2 had made an incorrect prediction, for if there was work to separate the charges, there would be emf and furthermore this work would become the potential difference which would make the charges move along the cable between zones A and B. However, of the 25 answers which answered student S2's comment, the great majority (N= 17) focused on it not being possible for emf to exist in a Van der Graaf generator.

The other two questions in the questionnaire were connected with the energy balance of an electrical circuit made up of a battery, cables and bulbs or resistances. Question Q4 is given below by way of example.

Q4. If a battery is connected to different circuits, what quantity does keep the potential difference between its terminals or the electromotive force of the battery?

Half the answers (N = 28) wrongly stated that what does not vary is the potential difference between the battery terminals. A small number of answers (N = 9; 16%) correctly reasoned that the electromotive force of a battery is a characteristic size thereof and that therefore it is kept constant. A quarter of the answers (N = 14) gave explanations unconnected with the theoretical framework or incoherent explanations.

Conclusions and Implications

Students' difficulties seem strongly linked to the absence of an analysis of the work carried out on the circuit and its energetic balance. In this regard, most 3rd course physics students still do not clearly understand the usefulness of concepts of potential difference and fem. The difficulties found and the contents of the analysis of the physics show the necessity to provide students with opportunities for them to reflect on the role played by the concepts of potential difference and electromotive force in the energetic model which explains the movement of the current in a simple electrical circuit. It will be necessary to design tasks and problems which lead to it being understood that the difference between the concepts of electromotive force and potential difference is given to measure different kinds of actions produced by radically different causes; the former due to non-conservative causes and the latter to conservative forces. This involves knowing that electromotive force is a quantity which quantifies a transfer of energy (from a battery to the charges of a circuit) associated with a non-conservative action.

References

Guisasola, J., Montero, A. y Fernández, M. (2005). Concepciones de futuros profesores de ciencias sobre un concepto 'olvidado' en la enseñanza de la electricidad: la fuerza electromotriz (Pre-service teachers' conceptions on a concept 'forgotten' in the teaching of electricity: electromotive force)*, Enseñanza de las Ciencias* 23(1), 47-60.

Guisasola, J., Furió, C. and Ceberio, M. (2008) Science Education based on developing guided research, Edited by M.V. Thomase in *Science Education in Focus* p. 55-85. Nova Science Publisher.

Mulhall P., Mckittrick B. & Gunstone R. 2001. A perspective on the resolution of confusions in the teaching of electricity, Research in Science Education 31, 575-587.

Watts, M, Gould, G. & Alsop, S., 1997. Questions of understanding: Categorising pupils' questions in Science. *School Science Review*, 79, 57-63.

ASSESSING ASSESSMENT TOOLS: TOWARDS QUESTIONNAIRE IMPROVEMENT THROUGH VALIDATION

Maria João Fonseca
Universidade do Porto

Patrício Costa, Leonor Lencastre
Universidade do Porto

Fernando Tavares
Universidade do Porto

Abstract

A major concern of Science Education researchers has been to assess the effectiveness of educational strategies, namely curricular structure, alternative teaching approaches and informal learning environments. In these studies, parameters such as knowledge, attitudes and behaviours are frequently addressed using diverse instruments chosen to answer precise research questions. In order to adopt corrective measures it is fundamental to obtain reliable results. If case-study approaches can be suitable to study in detail specific teaching activities, they hardly give a wider view of the students' universe. In contrast, questionnaires are commonly used assessment tools that allow broad surveys of multiple aspects among a vast universe of students. Furthermore, they provide considerable amounts of easily treatable information in relatively short periods of time. For these reasons, educational observatories, as the well-known EUROBAROMETER, tend to favour inquiry-based assessment. Regardless of their apparent simplicity, questionnaire-based instruments require proper procedures to provide reliable and valid data. To comprehensively assess the knowledge, interest and attitudes of elementary and high-school students from Porto towards biotechnology, and identify the sources they use and trust to gather information about it, an oriented questionnaire validation guideline is proposed, combining pilot work with statistical validation through exploratory factor scrutiny and reliability analysis.

Introduction

The ultimate goal of science education is the promotion of scientific literacy (Braun & Moses, 2004; Cabo Hernandez, Mirón, & Cortiñas Jurado, 2006), which is essential to assure that knowledgeable citizens are competent to take part in the democratic decision-making policies regarding scientific and technological processes. Biotechnology is one of the most controversial socio-scientific issues to date (Dawson, 2007; Klop & Severiens, 2007). Acknowledging the range and intensity of its impact on society, many studies have been conducted addressing public knowledge and attitudes towards biotechnology (Ping & Gutteling, 2009). However, the ones targeting student populations are still scarce (Dawson & Soames, 2006; Dawson, 2007; Firmino, 2007; Klop & Severiens, 2007; Prokop, Lešková, Kubitako, & Diran, 2007; Sáez, Niño, & Carretero, 2008, Uşak, Erdogan, Prokop, & Özel, 2009).

Quantitative inquiry surveys are the most suitable methodologies to achieve a broad characterization of a given population, especially when planning to establish correlations and make generalizations (Oppenheim, 1992, Black 1999). Nevertheless, questionnaires have limitations that are mainly related to their subjectivity (Black, 1999). Thus, it is necessary to design and administer these instruments following proper procedures that can assure the validity and reliability of the results they provide (Brace, 2008). Although validation should be an intrinsic element of large scale assessment, many studies fail to report these procedures (Petrić & Czarl, 2003). The decision to dismiss questionnaire validation may result from the perception that these statistical methods are not efficient given the time and effort they require, particularly since most research designs have strict timings (Dörnyei, 2002).

In this context, we designed a questionnaire to make the diagnostic of the knowledge, attitudes and decision-making capacity regarding biotechnology of elementary and high school students engaged in two different curricular formats: with and without Biology contents. The consistency of this questionnaire was improved through a pilot study and using statistical approaches to increase the data's reliability. Exploratory factor analysis and reliability analysis were conducted and their outcomes discussed.

Rationale

Questionnaire validation through a pilot study allows to identify, understand and address ambiguities that can interfere with the students' answers and compromise the reliability of the data gathered (Oppenheim, 1992, Black, 1999). It is important to assess the adequacy of the questionnaire's wording, length, structure and intelligibility, as well as the quality of the item's formulation, the scaling and the items' and questions' sequence (Black, 1999). The analysis of psychometric properties of the pilot study's data concerning the instrument is essential (Fabrigar, Wegener, MacCallum, & Strahan, 1999). Only then is the instrument ready to be administered.

Method

Participants

The questionnaire was applied in a representative sample of 92 students from three schools of Porto metropolitan area. Sample representativeness was assured by random selection of the participant schools. This sample comprised 46 students from two elementary 9th grade classes and 46 high-school students attending the 12th grade. 25 of these 12th graders were attending Biology and the other 21 were engaged in curricular formats that do not include Biology, namely economics, informatics, humanistic, and arts/design areas. Students from the 9th and 12th grades were chosen to participate in this study for two reasons. Firstly, these are the curricular years where more emphasis is given to biotechnology-related issues, according to the orientations of the Portuguese Ministry of Education (DEB, 2001; DGIDC, 2004). Secondly, since these are the concluding years of elementary and high-school, many of these students end their academic training without any other formal contact with biotechnology.

Measurement instrument

The questionnaire's pilot version was conceived considering instruments previously described (Miles, Ueland, & Frewer, 2003; Cabo Hernandez et al., 2006; Dawson & Soames, 2006; Gaskell et al., 2006; Dawson, 2007; Firmino, 2007; Prokop et al., 2007) and adapting relevant features to the Portuguese educational context. In addition, to better characterize the sample used and to increase the internal consistency of the instrument, new items were formulated. The questionnaire content was decided and assessed upon curricula and textbook analysis, thus assuring its validity. The pilot version consisted in a set of closed and semi-open questions, aiming to provide insights about the knowledge, attitudes, students' interest towards biotechnology and key sources of information used and trusted. Different scales were developed to assess each of these dimensions. The questions' wording, especially when translation was required, was carefully considered and the written language was adjusted to the characteristics of the respondents (Oppenheim, 1992). Negative phrasing was avoided and the items were formulated as objectively as possible to minimize bias (Black, 1999).

The questionnaire proposed consists in 14 questions, originally with a total of 65 items organized into 6 groups: *knowledge* (3 questions, 17 items), *attitudes* (2 questions, 18 items), *interest* (2 questions, 4 items), *comprehension of news about biotechnology* (1 question, 1 item), *sources of information about biotechnology used and trusted* (2 questions, 15 items) and *risk and benefit perception* (3 questions, 10 items). Despite its association with knowledge, the comprehension of news about biotechnology depends on how the information is divulged and therefore this item was isolated in its own category. The risk and benefit perception group comprises questions addressing behavioural intent regarding ethical and controversial issues, to assess the students' decision making capacity. Five factual data questions and one question concerning whether the students found the questionnaire interesting were added to better characterize the population and to enquire the receptivity towards this kind of approach, respectively. Five point Likert type scales were developed for each question, except for the questions in the *knowledge* section, that consists in a multiple choice question, a list of options and a True/ False/ Don't know question, and one question aimed at determining *sources of information used*, that presents a list of options to choose from. A *don't know* option was included in the True or False question to minimize social desirability bias (Black, 1999; Brace, 2008), and to assess the truthfulness of the students' answers.

Procedure

Data collection

The questionnaires were administered over a three month period, from October to December 2008, during class periods, without imposing time constraints, under the supervision of a teacher and/or the investigator. The time students took to complete the questionnaire was registered. Students were instructed to ask any questions resulting from difficulties in interpreting the questionnaire and regarding any words or concepts that might be unclear to them.

Data analysis

The data collected throughout the pilot study was codified according to a previously defined guideline, recorded and cleansed using the Statistical Package for the Social Sciences (SPSS) software version 17.0. After descriptive and missing values analysis of all the items in the questionnaire, the ones assessed by Likert type scales were subjected to an exploratory factor analysis (principal component analysis with varimax rotation) within their given dimension. Subsequently, by determining the Cronbach's alpha value for each factor identified, a reliability analysis was performed. The number of factors to retain following factor analysis was decided according to the Kaiser criterion (eigenvalues greater than 1) and the scree test (Hayton, Allen, & Scarpello 2004; Costello & Osborne, 2005). Items loading below 0.40, displaying low communality (below 0.40), cross-loading, freestanding or decreasing the scale's internal consistency were excluded from the analysis (Sharma, 1996; Fabrigar et al., 1999; Costello & Osborne, 2005; Hogarty, Hynes, Kromrey, Ferron, & Mumford., 2005). By comprehensively introducing some modifications in the pilot version, the final version of the questionnaire was obtained. It is worth to mention that although the items in the *knowledge* section and in the question regarding *sources of information used* were not subjected to factor analysis, both categories are included in the final version of the questionnaire. Item retention for these categories was based on the analysis of missing and ambiguous answers and on the results obtained for the scales subjected to factor analysis.

Results and Discussion

During the pilot study students did not identify major constraints that could justify any modifications regarding the content and structure of the questionnaire. Although there were three complaints about its length, all the students took less than 20 minutes to complete the whole set of questions. The maximum of missing values per item registered for the pilot sample ($n=92$) was 4,3%, so all the items were considered for subsequent analysis. The missing values were imputed using the series mean method in SPSS, given they were very limited and random (Huisman, 2000; Batista & Monard, 2003; Paul, Mason, McCaffrey, & Fox, 2008).

Factor analysis results, such as the ones summarized in Table 1, led to the introduction of some changes in the questionnaire's structure, and justified the need to be aware of aspects that influence the interpretation of the data gathered. The total number of items in the questionnaire was reduced by eliminating three items that appeared to be redundant: (i) *Genetically modified organisms can endanger the environment (answer True/False/Don't Know)* (ii) *Rate your agreement towards the following sentence - the ingestion of genetically modified foods has adverse effects on humans (1-I totally disagree to*

5-I totally agree); and (iii) *Rate your interest in participating in information campaigns about genetically modified organisms (1-I am not interested at all to 5-I am very interested).* It also became evident that a rearrangement of the items was necessary. The *attitudes* and the *risk and benefit perception* sections of the pilot version of the questionnaire were combined and their items re-structured according to the tri-partite model of attitudes (Klop & Severiens, 2007), in a section intended to assess the cognitive, affective and behavioural components of attitudes. Additionally, two items, one from each of these two initial sections, were combined into a new category, which factor analysis showed to be consistent with only one factor: *importance of biotechnology to the quality of life*. The sample's adequacy to factor analysis was confirmed by the Kaiser-Meyer-Olkin test results, except for the scales addressing the affective component of students' attitudes and the importance they give to biotechnology (KMO≤0,50). For these scales there seems to be a disperse pattern of correlations among variables, suggesting that a larger sample may yield better results (Sharma, 1996). Nevertheless, for each scale there is a significant ($p<0,05$) correlation between the variables tested, as it was demonstrated by the Bartlett's Test of Sphericity. Hence, it was decided to keep the factor structures identified for these scales throughout the main study to assess the effect of the sample's size on its adequacy. Hence, it was decided to keep the factor structures identified for these scales throughout the main study to assess the effect of the increase in the sample's size on its adequacy. The best factor structures identified by factor analysis for *attitude's cognitive component* and t*rust in sources of information*, exclude many of the items initially proposed for those scales (Table 2). However, failing to analyze such items would result in an important loss of information. Therefore, to improve the characterization of the student population, it was decided to include those items along with the factors identified for each scale in the final version of the questionnaire (Table 2).

From the reliability analysis carried out (Table 1), not all the Cronbach's alpha values obtained are sufficiently robust, scoring below 0.60 (Hair, Anderson, Tatham, & Black, 1998; Wasserman & Bracket, 2003), which may be due to the low number of items for the factors considered (Costello & Osborne, 2005). Although the number of items could be increased, the characteristics of the target population and the feedback obtained during the pilot study rendered a longer questionnaire inappropriate (Oppenheim, 1992; Dörnyei, 2002). At this point, it may be expected that the tendencies observed would be strengthened by increasing the sample size during the final study.

Exploratory factor analysis is an approach that requires many decisions and can produce misleading results (Fabrigar et al., 1999). Therefore, it is necessary to acknowledge that there are many factors affecting its outcomes, such as the design of the study, its aims and the data's properties (Costello & Osborne, 2005). In this study, our goal was to produce an adjusted instrument that can allow to reliably characterize the studied student population. Despite acknowledging that additional iterations would be beneficial to improve the consistency of the questionnaire's factor structures, it was decided to use the instrument in its current form during the main study. The validation process described proved to be of the utmost importance to emphasize aspects that were not obvious in the study's design.

Table 1. Scales' factor structure based on exploratory factor analysis (principal component analysis with varimax rotation) and reliability analysis. Coefficients below 0.30 were suppressed.

Scale	Factor Analysis							Reliability Analysis
	KMO	Identifiable factor	EV	% variance	Item	Communality	Item loading	Cronbach's alpha
Attitudes' Cognitive component	0,567	*Classical applications*	1,186	14,882	- *Use of yeast in the production of bread, wine and beer*	0,691	0,802	0,639
					- *Use of yeast in animal food production*	0,761	0,868	
		Agro-food applications	1,507	18,842	- *Improvement of the growth of plant in saline environments by altering their genes*	0,714	0,829	0,623
					- *Production of pesticide resistant plants by gene manipulation*	0,638	0,670	
					- *Genetic modification of tomatoes to make them ripen more slowly and have a longer shelf life*	0,527	0,621	
		Bio-medical applications	2,518	31,472	- *Utilization of genetically modified cows for the production of medicines for human*	0,714	0,819	0,665
					- *Use of insulin produced by bacteria*	0,587	0,730	
					- *Organ transplant from transgenic animals to humans*	0,578	0,717	
Attitudes' Affective component	0,449	*Embryonic cell utilization*	1,313	32,836	- *It is our duty to authorize investigation that may lead to the development of more efficient medical treatments, even if it implies using embryonic stem cells*	0,770	0,877	0,367
					- *It is wrong to use embryonic stem cells in biomedical research, even if it may contribute to the development of medical treatments*	0,688	0,487	
		Control capacity	1,171	29,267	- *The labels of transgenic food should specify weather the food or any of its ingredients is genetically modified*	0,623	0,658	0,333
					- *Each of us is capable of determining our intake of transgenic foods*	0,404	0,635	
Attitudes' Behavioural component	0,624	*Buying intent*	2,377	36,618	- *Buy transgenic foods if they were available in supermarkets*	0,724	0,850	0,633
					- *Buy medicines obtained by genetically manipulation*	0,424	0,621	
					- *Buy transgenic foods if they were healthier than other foods*	0,398	0,532	
					- *Buy transgenic foods if they were less expensive than other foods*	0,447	0,655	
		Access to genetic information	1,118	18,637	- *Do a genetic test for medical diagnostic*	0,805	0,894	0,693
					- *Give the police access to your genetic information*	0,697	0,806	
Interest	0,674	*Interest*	2,375	59,373	- *Rate your interest towards biotechnology*	0,367	0,606	0,768
					- *Listen to news about biotechnology*	0,575	0,758	
					- *Read articles or watch TV shows about biotechnology*	0,821	0,906	
					- *Search the web for subjects related to biotechnology*	0,612	0,782	
Trust in sources of information	0,661	*Industry*	2,106	42,113	- *Pharmaceutical industry*	0,610	0,772	0,630
					- *Agro-food industry*	0,673	0,819	
					- *Health industry*	0,576	0,711	
		NGO's	1,210	24,193	- *Environmental organizations*	0,743	0,862	0,671
					- *Consumer rights organizations*	0,714	0,821	
Importance	0,500	*Importance*	1,395	69,740	- *How important do you think biotechnology is to the quality of life?*	0,697	0,835	0,563
					- *Do you agree that future generations will benefit from biotechnology's medical applications?*	0,697	0,835	

Table 2. Questionnaire structure following validation. The factor structure identified for each scale is highlighted in bold.

Section / Scale	Example of questions used
Factual questions	School you are attending; Grade you are attending; Age; Sex; Course you are attending
Knowledge assessment	Q1: Biotechnology can be defined as a set of processes… (Select the option you most agree with) (i)… in which recombinant DNA technology is used; (ii)… applied to investigation and product development; (iii)… which involves cell and tissue culture; (iv)… by which genetically modified organisms (GMOs) can be developed. Q2: From the biotechnological applications listed, select the one(s) you know. Production of medicines and vaccines; Production of hormones; Production of organic products, such as milk or yogurt; Recovery of contaminated soils using genetically modified bacteria; Waste treatment; Production of insect and pesticide resistant plants; Utilization of plants with industrial purposes, namely the production of cosmetics, plastics and fuels; Production of amminoacids and vitamins. Q3: Answer the following questions using True/False/I don't know. The ingestion of genetically modified foods can induce gene alterations; Cloning and genetic engineering are identical processes; It is impossible to transfer genes from plants to animals; Genetically modified organisms contain dangerous chemicals; …
Importance assessment	Q4: **How important do you think biotechnology is to the quality of life? (1-Not at all important to 5-Very important)[a]** Q6c: **Rate your agreement with the following sentence (1-I disagree completely to 5-I agree completely): Do you agree that future generations will benefit from biotechnology's medical applications?[a]**
Attitudes assessment – cognitive component	Q5: Rate your approval towards the following activities (1-I do not approve it at all to 5-I approve it completely). **Use of yeast in the production of bread, wine and beer[b]**; **Use of yeast in animal food production[b]**; Use of genetically modified organisms in waste treatment; **Improvement of the growth of plant in saline environments by altering their genes[c]**; Treatment of genetic disorders by embryonic gene manipulation; Treatment of genetic disorders by human gene manipulation; Insertion of plant genes into animals; **Utilization of genetically modified cows for the production of medicines for humans[d]; Production of pesticide resistant plants by gene manipulation[c]**; **Genetic modification of tomatoes to make them ripen more slowly and have a longer shelf life[c]**; **Use of insulin produced by bacteria[d]**; **Organ transplant from transgenic animals to humans[d]**; Medical treatments through human cloning.
Attitudes assessment – affective component	Q6: Rate your agreement with the following sentences (1-I totally disagree to 5-I totally agree). **It is our duty to authorize investigation that may lead to the development of more efficient medical treatments, even if it implies using embryonic stem cells[e]**; **The labels of transgenic food should specify weather the food or any of its ingredients is genetically modified[f]**; **It is wrong to use embryonic stem cells in biomedical research, even if it may contribute to the development of medical treatments[e]**; **Each of us is capable of determining our intake of transgenic foods[f]**.
Attitudes assessment – behavioural component	Q7: How often would you… (1-Never to 5-Always) **Buy transgenic foods if they were available in supermarkets[g]**; **Buy medicines obtained by genetically manipulation[g]**. Q13: How often would you… (1-Never to 5-Always) **Do a genetic test for medical diagnostic[h]**; **Give the police access to your genetic information[h]**; **Buy transgenic foods if they were healthier than other foods[g]**; **Buy transgenic foods if they were less expensive than other foods[g]**.
Interest assessment	Q8: **Rate your interest towards biotechnology (1-I am not interested at all to 5-I am very interested)[i]**. Q9: How often do you… (1-Never to 5-Many times) …**listen to news about biotechnology[i]**; …**read articles or watch TV shows about technology[i]**; …**search the web for subjects related to biotechnology[i]**.
Comprehension of news assessment	Q10: Rate your difficulty in understanding news about biotechnology (1-very low to 5-Very high).
Assessment of sources of information used	Q11: From which of the following sources do you most commonly obtain information about biotechnology? TV; Radio; Newspapers, Magazines; Scientific magazines; Internet; Textbooks; Teachers; Friends; Family; Others.
Assessment of trust in sources of information	Q12: Rate your trust in the following sources of information about biotechnology (1-I do not trust it/them at all to 5-I trust it/them completely): Media; Scientific magazines; **Pharmaceutical industry[j]**; **Agro-food industry[j]**; **Health industry[j]**; Governmental agencies; Universities; Scientists; Internet; **Environmental organizations[k]**; **Consumer rights organizations[k]**; European Community; Medical doctors; Politicians.

Note. Items identified with the same letter (a,b,c,d,e,f,g,h,i,j,k) contribute to the same factor.

Conclusions and Implications

We believe that the quality of the data gathered through an inquiry-based methodology can be improved by an adequate validation of the instruments used. In addition, this procedure may be fundamental to assess and optimize the reliability and validity of those instruments and the results they provide. When developing new questionnaires, analysis of psychometric features can help reduce bias introduced by the author's own expectations towards the students' answers. Furthermore, it can unveil intrinsic and unpredicted conditioning factors. Validation also becomes crucial when planning to use already existent instruments, since the population to be assessed presumably differs from the ones for which those instruments were originally intended.

Overall, the work carried out underlines a guideline to similar diagnostic studies. A validation strategy was tested and its outcomes discussed, demonstrating that the investment in these time consuming procedures can improve the quality of the data gathered through quantitative assessment methodologies.

Acknowledgements

The authors are grateful to all the participant students and teachers and to Catarina L. Santos for helpful comments and suggestions on the manuscript. Maria João Fonseca is supported by the FCT fellowship SFRH/BD/37389/2007.

References

Batista, G., & Monard, M. (2003). An analysis of four missing data treatment methods for supervised learning. *Applied Artificial Intelligence*, 17(5), 519-533.

Black, T. R. (1999). *Doing quantitative research in the Social Sciences. An integrated approach to research design, measurement and statistics.* London: SAGE Publications.

Brace, I. (2008). *Questionnaire Design. How to Plan, Structure and Write Survey Material for Effective Market Research.* London: Kogan Page Publishers.

Braun, R. & Moses, V. (2004). A public policy on biotechnology education: what might be relevant and effective? *Current Opinion in Biotechnology*, 15, 246-249.

Cabo Hernández, J. M., Mirón, E. & Cortiñas Jurado, J. R. (2006). Opiniones e intenciones del professorado sobre la participación social en ciencia y tecnologia. El caso de la biotecnologia. *Revista Eureka Enseñanza e Divulgación de Ciencia*, 3(3), 349-369.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment, Research and Evaluation*, 10(7). Retrieved from http://pareonline.net/getvn.asp?v=10&n=7

Dawson, V. (2007). An Exploration of high school (12-17 year old) students' understandings of, and attitudes towards biotechnology processes. *Research in Science Education*, 37(1), 59-73.

Dawson, V. & Soames, C. (2006). The effect of biotechnology education on Australian high school student's understandings and attitudes about biotechnological processes. *Research in Science and Technological Education*, 24(2), 183-198.

DEB. (2001). *Orientações curriculares do Ensino Básico - Ciências Físicas e Naturais*. Lisboa, Portugal: Ministério da Educação.

DGIDC. (2004). *Programa de Biologia 12º ano. Curso Científico - Humanístico de Ciências e Tecnologias*. Lisboa, Portugal: Ministério da Educação.

Dörnyei, Z. (2002). *Questionnaires in second language research: Construction, administration, and processing.* New Jersey: Lawrence Erlbaum Associates.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 3, 272-299.

Firmino, M. (2007). Biotecnologia – estudo exploratório das percepções e atitudes de professores e alunos. (Masters dissertation, Faculdade de Ciências, Universidade do Porto).

Gaskell, G., Stares, S., Allansdottir, A., Allum, N., Corchero, C., Fischler, C., . . . Wagner, W. (2006). *Europeans and biotechnology in 2005: Patterns and trends. Final report on Eurobarometer 64.3*. Brussels: European Commission's Directorate-General for Research.

Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis.* (5th ed.). New Jersey: Prentice-Hall.

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: a tutorial on parallel analysis. *Organizational Research Methods*, 7(2), 191-205.

Hogarty, K. Y., Hynes, C. V., Kromrey, J. D., Ferron, J. M., & Mumford, K. R. (2005). The quality of factor solutions in exploratory factor analysis: the influence of sample size, communality and overdetermination. *Educational and Psychological Measurement*, 65(2), 202-226.

Huisman, M. (2000). Imputation of missing item responses: some simple techniques. *Quality and Quantity*, 34(4), 331-251.

Klop, T. & Severiens, S. (2007). An exploratory of attitudes towards modern biotechnology: a study among Dutch secondary school students. *International Journal of Science Education*, 29(5), 663-979.

Miles, S., Ueland, Ø. & Frewer, L. J. (2003). Public attitudes towards genetically-modified food. *British Food Journal*, 10(4), 246-262.

Oppenheim, A. N. (1992). Questionnaire design, interviewing and attitude measurement (new ed.). London; New York: Continuum International Publishing Group.

Paul, C., Mason, W. M., McCaffrey, D., & Fox, S.A. (2008). A cautionary case study of approaches to the treatment of missing data. *Statistical Methods and Applications*, 17, 351-372.

Petrić, B, & Czárl, B. (2003). Validating a writing strategy questionnaire. *System*, 31, 187-215.

Ping, R. R. & Gutteling, J.M. (2009). The development of public perception research in the genomics field. An empirical analysis of the literature on the field. *Science Communication*, 31(1), 57-83.

Prokop, P., Lešková, A., Kubitako, M. & Diran, C. (2007). Slovakian students' knowledge of and attitudes toward biotechnology. *International Journal of Science Education*, 29(7), 895-907.

Sáez, M. J., Niño, A. G. & Carretero, A. (2008). Matching society values: students' views of biotechnology. *International Journal of Science Education*, 30(2), 167-183.

Sharma, S. (1996). *Applied multivariate techniques*. USA: Jonh Wiley & Sons, Inc.

Uşak, M., Erdogan, M., Prokop, P., & Özel, M. (2009). High school and university students' knowledge and attitudes regarding biotechnology. *Biochemistry and Molecular Biology Education*, 37(2), 123-130.

Wasserman, J. D., & Bracken, B.A. (2003). Psychometric characteristics of assessment procedures. In J. R. Graham, & J. A. Naglieri (Eds.), *Handbook of Psychology: Assessment psychology* (Vol. 10). (pp. 43-66). Hoboken, N J: John Wiley & Sons, Inc.

DIGITAL PORTFOLIO ASSESSMENT OF SECONDARY STUDENTS' SCIENTIFIC ENQUIRY SKILLS: THE E-SCAPE PROJECT

Dan Davies
Bath Spa University

Abstract

This paper reports on the development of a digital system to assess secondary school students' scientific enquiry skills as an alternative to current modes of assessment at GCSE-level (age 16) in England. Part of the larger 'e-scape' project comprising parallel developments in design & technology and geography, the research has focussed upon the development of a controlled assessment activity undertaken by students in a three-hour period, using hand-held digital devices to record their responses using text, digital photographs, drawings, spreadsheets, video and audio commentary. Following trialling with 135 students in 3 schools, a team of 6 judges were invited to assess the resulting electronic portfolios using a 'Thurstone pairs' approach (Greatorex et al. 2008), yielding a rank order with an acceptably low error value and moderate correlation with predicted GCSE grades for the students concerned. Moreover, judges, teachers involved in trialling and students viewd the activity and assessment system as more motivating and potentially more valid in capturing scientific capability than existing modes of practical assessment in use in England. Awarding bodies (examination boards) are currently evaluating this development work with a view to introducing alternative forms of 'e-portfolio' assessment in the future.

Introduction

Much political and educational debate in the UK has been focussed on the challenge of coursework assessment at GCSE-level (age 16) (Roberts & Gott 2006, Qualifications and Curriculum Authority (QCA) 2005). The educational value of coursework is widely understood and appreciated, particularly in terms of the process-skills that it is able to probe and the opportunities for student creativity (Bullock et al. 2002). But increasingly it is seen as providing potentially invalid and unreliable data on the performance of students (QCA 2005). These potential inaccuracies arise from several sources:

• the activities on which they are based are not standardised;
• the circumstances in which they are conducted are not standardised;
• the time allocations for the activity are not standardised;
• the resources available to learners are not standardised;
• the teacher-support for learners is not standardised;
• the authorship of out-of-school work cannot be guaranteed (Kimbell et al. 2007).

The above activity-management issues are then exacerbated by other problems at the assessment end of the process. For high-stakes assessment, the lack of evident equity in the administration of coursework, allied to concerns over the reliability of the assessment has led, in the case of science, to GCSE syllabi and assessment criteria being extensively reviewed and rewritten (Qualifications and Curriculum Authority (QCA) 2006). Coursework was replaced by Investigative Skills Assignments (ISAs) (AQA 2008) – essentially practical examinations consisting of an experiment undertaken by the whole class in groups during one lesson and a 45 minute written 'evidence test' (Roberts & Gott 2006).

Rationale

In response to some of the above concerns in relation to the portfolio assessment of GCSE design & technology, the UK government Qualification and Curriculum Authority (QCA) asked the Technology Education Research Unit at Goldsmiths' University of London to undertake research to examine the extent to which - and the ways in which - innovation and team-work might be more fully recognised and rewarded in assessment processes. The principal outcome of that project was a developed, paper-based, portfolio assessment system which served as a hybrid between a formal examination and a form of coursework. The consensus of teachers and learners involved in the project was that the portfolio system acted as a dynamic force to drive the activity forward with pace and purpose (Kimbell et al. 2007). The emerging field of 'e-learning', with its potential to support assessment for learning through digital systems, led TERU to develop a proposal for a digital approach to coursework portfolio assessment, called project 'e-scape' (e-solutions for creative assessment in portfolio environments). The project was designed in three closely interlinked phases: 1) proof of concept, 2) working prototype, 3) proof of transferability and scalability. During phase 3 (June 2007 – December 2008) the development work undertaken in D&T was extended to two other areas of the curriculum: geography and science (Kimbell *et al.* 2009). As part of this phase of the research, the author was contracted, along with a secondary science consultant based at Sheffield Hallam University (UK) to develop and trial e-scape assessment tasks in science as an alternative to ISAs. This paper reports on the process and outcomes of the science element of phase 3.

Research Questions

The essential question that phase 3 of the e-scape project sought to answer concerned the transferability of the e-portfolio model of assessment that had been developed in the context of design & technology GCSE coursework. Accordingly, our research question was as follows:

> To what extent is the 'e-scape' model for dynamic e-portfolio assessment of D&T capability transferrable to a GCSE science assessment context?

This question was broken down into three specific sub-questions, each addressing a particular stage in the development process:

a) Can the 'e-scape' model facilitate the development of purposeful, active science learning activities?

b) Can such activities generate evidence for valid assessment of pupils' scientific capabilities?

c) To what extent can the resulting pupil e-portfolios be judged using a 'Thurstone pairs' approach (Greatorex et al. 2008) to produce a reliable rank order for future grade awards?

Methods

Since the researchers in this study were also the task developers, the methodology adopted was broadly one of action research. In order to provide evidence against the above research questions, data were collected from the following sources:

- Observations of prototype assessment activities being trialled (13 trials over 10 months in 4 schools involving 135 pupils aged 14-15)
- Questionnaires and group interviews with pupils immediately following trials (n = 135)
- Interviews with teachers involved in trials (n = 5)
- Interviews with judges recruited to assess e-portfolios (n = 5)
- Analysis of results of e-portfolio judging process and subsequent mock grade awarding meeting.

We collected the above data during each of the four sub-phases of the task development and trialling process, reflecting upon the outcomes of each trial in order to inform subsequent development as consistent with an action research methodology. The four sub-phases were as follows:

1. Development and piloting of assessment activities in a paper-based format (June – December 2007). Working closely with heads of science in three secondary schools, the team developed a range of contexts and sub-tasks which were piloted with groups of students and sequenced together into assessment activities. At the end of this phase, one such activity (set in the context of road safety) was selected for further development.
2. Transfer of the selected assessment activity to the e-scape electronic examination management system (EMS) (January to March 2008). This consists of a web-based digital authoring tool in which the activity is constructed from a series of sub-tasks; a portable server and wireless network used to control the activity in the classroom; and a class set of handheld computers ('Ameos') wirelessly linked to the server so students can both see the sub-tasks in sequence and respond using a range of different tools – text, drawing, data entry in spreadsheets, video and still photography and voice recording. Following further classroom piloting using EMS, the selected assessment activity was further refined.
3. Trialling of the standardised assessment activity (3 hours duration) with Y10 students (aged 14-15) under controlled conditions in three schools (April-July 2008). This resulted in the generation of 55 e-portfolios of assessment evidence on the portable server, which were then uploaded into the web-based judging system.
4. Judging the e-portfolios using a mixture of 'Thurstone pairs' and rank ordering (Greatorex et al. 2008) (September – December 2008). Six judges experienced in GCSE science assessment were recruited and trained, then asked to judge 22 portfolios presented to them in pairs by the e-scape online judging system. Once 220 pair comparison judgements had been made (37 per judge) the system generated a rank order of e-portfolios from the trials, which was then used as the basis of a mock awarding meeting involving a senior examiner from an examination board.

Qualitative data from observations and interviews were analysed against a framework derived from the above research questions and the following list of criteria for the development of assessment tasks derived from previous phases of the e-scape project:

- Tasks should reflect the very best active-learning models within science
- The duration of tasks is determined by the need for authenticity
- Tasks should be seen as valid activity by the science community
- Tasks should encourage elements of group/team contribution in support of individual performance
- Tasks should be based around an overriding task/purpose within a science context
- Tasks should be thoroughly contextualised for learners, providing meaning and purpose for the activity
- Tasks should create standardised demands on learners but encourage diverse responses from them
- Tasks should unfold through a structured and interconnected series of sub-tasks of varying length
- The encouragement of individual creative responses will need to be signalled and supported
- Tasks should be resourced with whatever equipment and resources are necessary for authentic activity
- Learners' work should be undertaken on hand-held digital tools using the e-scape interface.

Quantitative data concerning student performance generated during the online pairs-judging process were analysed statistically using Rasch analysis of the judgement data (Pollitt and Crisp 2004) to generate estimates of the 'ability' of each student (or the 'quality' of each script), with standard errors and misfit statistics.

Results

Observations of students undertaking the trial version of the e-scape science assessment task (sub-phase 3 above) suggest that it required them to:

- explore the science of everyday situations (a handling collection of road-safety related materials and two short video presentations on pedestrian safety);
- interpret official data on vehicle stopping distances and the effectiveness of various traffic-calming measures;
- plan an experiment to investigate the relationship between vehicle speed and impact force;
- make predictions using sketch graphs;
- use digital timing and force measurement equipment, making judgements on accuracy, precision and reliability;
- collect, compare and interpret experimental data using mathematical models;
- comment on anomalous data and question the validity of official statistics.

Summarising the pupil questionnaires and group interviews, there was an overwhelming response that the e-scape science assessment task had been a 'fun and engaging activity', though they had found some bits quite difficult. A majority agreed that e-scape tasks were a good way to assess their ability to think and act like scientists. Unsurprisingly there was considerable enthusiasm for the hand-held technology used and the e-scape system for 'harvesting' e-portfoliios in real-time, though several students expressed frustration at the technical issues they encountered, such as 'freezing' of the devices, necessitating re-setting and consequent time lost in the activity. A few went as far to suggest that the technology had 'got in the way' of being able to show how good they were at science. Similarly, some students expressed frustration with the rigid timing of sub-tasks, which forced them to move on to the next part of the activity before they were ready.

Interviews with teachers following the assessment task trials elicited the following comparisons between the e-scape tasks and the ISAs which the teachers were currently using to assess students' scientific investigative skills:

- e-scape assessment task was more engaging for learners than ISAs;
- longer duration and a more sustained activity than ISAs;
- a more challenging, sustained task than ISAs, broken down into manageable sub-tasks, moving through cycles of action and reflection;
- innovative use of video clips and handling collections to strengthen the context and stimulate creative thought in the pupils;
- an emphasis upon 'extracting the science' from everyday situations and bringing the evidence from investigation to bear on a 'real world problem', by contrast with decontextualised nature of ISAs;
- requiring longer answers than ISAs and providing more opportunities for explanation
- offering access through voice recording and video for those who struggle with written work ;
- a higher degree of pupil autonomy than ISAs in designing investigations to test hypotheses;
- formal recognition of peer input to candidates' ideas and experimental designs, found to be valued by secondary school students (Cowie 2005);
- a similar mix to ISAs between collaborative group work and individual reflection, with clear boundaries between these phases.

Overall, teachers' responses to the trialled road safety activity were that it represented effective science pedagogy and an opportunity for students to engage in 'real science' (research sub-question a above). However, they were concerned by the amount of equipment used and felt that the demands on the administrator were very high. They also felt that ISAs were potentially quicker and easier to assess, whilst e-scape tasks required holistic judgement which was perceived as difficult and impressionistic. By comparison, those who had actually been involved in the judging process – though the idea of pairs comparison had felt strange at first - agreed in retrospect that it had worked well and had been surprisingly easy. Judges were enthusiastic about using judging as part of general science assessment practice and felt that students could potentially improve their own performance by being involved in the judging process. Although the judges regarded overall student performance evidenced by the e-scape e-portfolios to be somewhat lower than might be expected from a GCSE examination, they acknowledged that the trial groups had been at least a year younger than students normally taking GCSE science; they had received no coaching for the test; the technology was unfamiliar and the conceptual and procedural demands of the e-scape task were felt to be challenging.

At the end of the Rasch analysis of the quantitative pupil performance data undertaken by the online e-scape 'pairs engine' the rank order of the 22 portfolios in the judging process, together with the parmeter value error plot generated is shown in figure 1. In this figure, each point represents a single portfolio, with the lowest ranking (that which 'lost' all pairs comparisons with other student's work) on the far left, and the highest ranking (that which 'won' all pairs comparisons) on the far right. Error bars represent the difference between judgements made. As can be seen from figure 1, the errors on judging 19 of the 22 portfolios were acceptably small (around 0.5 of a parameter), suggesting a high level of reliability in relation to research sub-question c above. Only the top and bottom portfolios (ranks 1 and 22) produced large errors, which is a feature of the system rather than the reliability of marking – since rank 1 consistently 'won' all its pairs comparisons with all judges, the system could not tell how much 'better' it was than it's nearest rival. This relatively high reliability is reflected in other studies of the 'Thurstone pairs' method (Kimbell et al. 2007, Greatorex et al. 2008).

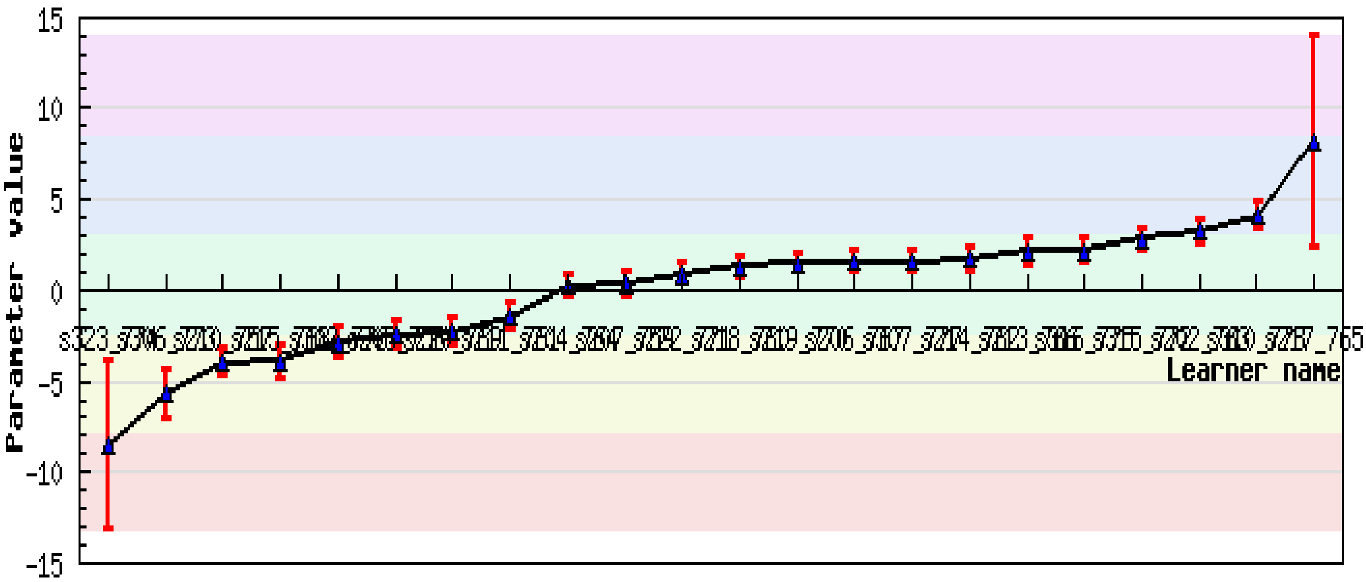


Figure 1. Outcome of pairs ranking process and parameter value error plot.

In order to assign nominal 'grades' to the students' work , a team of two researchers reviewed the online judging history for each student portfolio in the rank order, analysing the notes each judge had made on each portfolio, the degree of agreement between judges and the identification of key discriminating criteria, which were identified as follows:

1. General level of scientific explanation
2. Interpretation of data/ choice of graph
3. Understanding of technical terms (e.g. accuracy, precision)
4. Choice of independent variable for their own investigation
5. Perseverance in collecting data
6. Evidence of reasoning with data (e.g. identification of anomalies, reflection upon reliability of test)
7. Engagement with equation
8. Ideas for testing road safety 'solutions'.

In order to generate descriptors for different levels of performance ('grades') we decided to summarise the judges' comments on each portfolio in terms of performance against the above criteria. Each we rated on a scale of zero (no engagement, data missing) to 3 (good), then aggregated the scores for each portfolio in the rank order using this atomised approach (see table below). The Spearman (ordinal) correlation between these aggregated scores and the rank order was high at 0.964, though it should be borne in mind that we worked down the rank order in this exercise. The aggregated scores appeared to fall into five levels of performance, to which we gave the 'tentative' grades B to F, since in our opinion no portfolio provided evidence of 'A' or 'A*' performance in GCSE terms, and the lowest two in the rank order were very weak. We next compared these tentative grades against the statutory science test results from the previous year (Y9 SATS) and the predicted GCSE grades for the students concerned. The Pearson (interval by interval) correlations between the data emerging from the judging and grade awarding process and data from the schools are given in table 1:

Table 1. Correlations between e-scape judging process and existing forms of science assessment

Between	And	Pearson's R	Significance
Rank order	SATS score	0.507	0.02
Rank order	Predicted GCSE grade	0.537	0.02
E-scape tentative grade	SATS score	0.431	0.06
E-scape tentative grade	Predicted GCSE grade	0.467	0.04

As can be seen from table 1 above, the correlations between the rank order generated by the pairs comparison process and the school SATS and predicted GCSE grades were slightly stronger (and more significant) than those between the tentative grades we had assigned them during the awarding process. Associations are generally modest (though generally significant at the 5% 0.05 level).

Conclusions and Implications

There is some evidence from the development and trialling of a controlled digital assessment activity and from the judgement of the resulting e-portfolios, that the e-scape system is capable of generating an alternative to current models of GCSE science assessments of student's procedural knowledge and investigative skills. In relation to the research question, this phase 3 project has demonstrated the transferability of the e-scape assessment system from design & technology to science. Although there are significant differences in approach (for example a three-hour activity rather than the six-hour D&T model) the science activity has retained many of the distinctive e-scape features, such as a strong context; the use of handling collection and other stimulus resources to engage learners with the activity; and the sense of a strong purpose and end-point. There is evidence that this activity is motivating for students, looks like 'real science' to teachers and is sufficiently discriminating to allow a reasonable degree of reliability in pairs judging.

However, a number of issues have arisen concerning task development, the technology and the judging process. Generally, the e-portfolios generated by students during the activity did not offer a full picture of student capability; they ranged from poor to average whilst high attainers were not able to show their full potential. This could be related to the difficulty of some of the questions, the light-hearted attitudes of some students during the pilot and/or technical issues with timing and data transfer. Interestingly, some of the judges felt that the original paper-based portfolios offered more evidence of higher attainment than the electronic versions. Although there was a range of attainment in the portfolios, this was relatively narrow and many judgements were hard to make as the questions were in many cases not discriminating (all students tended to struggle with some of them). Many boxes were incomplete, or data was missing. Some of the technical 'glitches' may have given a false impression – e.g. incomplete and missing answers could have been due to technical problems rather than lack of knowledge.

e-scape science requires further development for use as a curriculum activity which could prepare learners for its use in assessment, and as a rigorous, fair test of scientific attainment:

- A range of contexts and activities spanning the breadth of the science curriculum
- Shorter activities that could fit within a single lesson or sequence of lessons
- Activities which use each of the features of the hand-held devices separately (text, picture, video, voice, spreadsheet) to enable students to become familiar and confident with these
- Access to a wider range of web-based and other teaching resources within each activity (e.g examples/explanations of accuracy and reliability)
- More choice given to students in terms of the communication tools they could choose for each box and the design of investigations.

These are only preliminary findings from the trial of one enquiry activity; it should be remembered that process skills in science tend to be context-dependent and that different investigations make different procedural and conceptual demands (Matthews & McKenna 2005). However, the holistic nature of the judgements made through the pairs comparison process reflect the holistic nature of the scientific enquiry process, which existing modes of assessment may fail to capture by breaking it down into discrete skills (Matthews & McKenna 2005). Further research is needed to more fully evaluate the 'affordances' given to students by the use of hand-held devices in the assessment process, particularly the multi-modality offered by the use of drawing, text, voice recording, digital photography and video.

References

Assessment and Qualifications Alliance (AQA) (2008), *General Certificate of Secondary Education: Science A 4461, 2008*. Retrieved from: http://www.aqa.org.uk/qual/pdf/AQA-4461-W-SP-08.PDF.

Bullock, K., Bishop, K., Martin, S. & Reid, A. (2002) Learning from Coursework in English and Geography. *Cambridge Journal of Education, 32*(3), 325-340.

Cowie, B. (2005) Student commentary on classroom assessment in science: a sociocultural interpretation. *International Journal of Science Education, 27*(2), 199-214.

Greatorex, J., Novakovic, N. & Suto, I. (2008) What attracts judges' attention? A comparison of three grading methods, *Annual Conference of the International Association for Educational Assessment*, Cambridge.

Kimbell, R., Wheeler, A., Miller, S. & Pollitt, A. (2007) *E-scape portfolio assessment phase 2 report*. London: Department of Design, Goldsmiths, University of London.

Kimbell, R., Wheeler, A., Miller, S. & Pollitt, A. (2009), *E-scape portfolio assessment phase 3 report*, Department of Design, Goldsmiths, University of London. Available on-line: www.gold.ac.uk/teru/projectinfo/projecttitle,5882,en.php (accessed 15.9.09)

Matthews, P.S.C. & McKenna, P.J. (2005) Assessment of Practical Work in Ireland: A critique, *International Journal of Science Education, 27*(10), 1211-1224.

Pollitt, A. and Crisp, V. (2004) Could Comparative Judgements Of Script Quality Replace Traditional Marking And Improve The Validity Of Exam Questions? Retrieved from: http://www.gold.ac.uk/media/113798_Could_comparative_judgements.pdf

Qualifications and Curriculum Authority (QCA) (2006) *GCSE Criteria for Science*. London: QCA.

Qualifications and Curriculum Authority (QCA) (2005) *A review of GCE and GCSE coursework arrangements*. London: QCA.

Roberts, R. & Gott, R. (2006) Assessment of performance in practical science and pupil attributes. *Assessment in Education: Principles, Policy & Practice, 13*(1), 45-67.

PREDICTIVE VARIABLES FOR ITEM GENERATION IN SCIENCE RELATED TO LINEAR FUNCTIONS[1]

Martin Riopel
Université du Québec à Montréal

Fadia Sakr, Gilles Raîche, Patrice Potvin, Patrick Charland & Valérie Djédjé
Université du Québec à Montréal

Abstract

Item generation is a research effort to develop general models to predict evaluation tasks parameters such as difficulty by using other fundamental characteristics of the tasks. Once validated, these models could then be used to automatically generate adaptive new tasks (or items) to evaluate students. This paper presents the results of a research on predictive variables about evaluation tasks in science related to linear functions for secondary level students. More precisely, the research has studied 100 items produced by the Société de Gestion du Réseau Informatique des Commissions Scolaires (GRICS) that were administered to 6910 students in 22 schools of the Quebec province in the eastern part of Canada between 1996 and 2003. The proposed model is based on eight predictive variables and can predict success rate with a correlation coefficient of 0.78. The model has been explicitly designed to be usable for classification of existing items but also for generation of new items. As a proof of concept, the proposed model has then been used to generate about 3 000 nonequivalent items with predicted success rates ranging from 0% (very difficult task) to 84% (very easy task). It could be used online by teachers or researchers to generate evaluation tasks or by computerized systems for adaptive evaluation.

Introduction

Item generation is a research effort to develop general models to predict evaluation tasks parameters such as difficulty by using other fundamental characteristics of the tasks. Once validated, these models could then be used to automatically generate adaptive new tasks (or items) to evaluate students. As described in Riopel et al. (2009), the first research on item generation can be traced back to the sixties when Hively, Patterson and Page (1968) proposed evaluation tasks constructed from a combination of fixed and variable elements. This approach is still useful today in new areas of research like assessment engineering (Luecht, 2007). Bejar, Morley and Bennett (2003) described two main types of research about item generation based on weak theory and strong theory approaches. Strong theory approach generates and validates directly new items based on explicit and a priori cognitive models. Weak theory approach generally prefers to study existing items in order to construct and to validate a predictive model a posteriori. As presented in Irvine and Kyllonen (2002), most of the research on item generation concerned general cognitive tasks not specifically related to science proficiency. The research presented in this paper uses the weak theory approach to identify predictive variables about evaluation tasks in science related to linear functions for secondary level students.

[1] This research has been supported by grant #410-2007-1357 from the *Social Sciences and Humanities Research Concil of Canada* (SSHRC) and grant #119232 from the *Fonds québécois de la recherche sur la société et la culture* (FQRSC). The authors gratefully acknowledge the assistance of Diane Leduc.

Rationale

Producing good assessment tasks in authentic context is not an easy thing to do. In schools, teachers don't always take enough time to do so and the quality of assessment could be improved. Meanwhile, sets of good tasks are produced by researchers interested in comparing achievement in different populations. Two well known examples are the Trends in International Mathematics and Science Study (TIMSS) and the Programme for International Student Assessment (PISA). The good tasks created for these studies are usually not released to teachers for obvious reasons: once released they can be no longer used in further research projects. One way to avoid this problem could be, instead of using research effort only to create the tasks themselves, to produce and validate also general models for assessment tasks. This way, different but equivalent tasks could be released to teachers to be used on a daily basis as long as these general models could produce really a lot of different tasks. This paper presents the first step of a project that aims to produce such a general model for assessment tasks in science. Since linear relations are used everywhere in sciences, they seemed like a logical a starting point. The main research question is: *what are the predictive variables for the difficulty of assessment tasks involving linear relations?*

Methods

The research studied 100 items produced by the Société de Gestion du Réseau Informatique des Commissions Scolaires (GRICS) that were administered to 6910 students in 22 schools of the Quebec province in the eastern part of Canada between 1996 and 2003. The proposed model is based on the eight most effective predictive variables (from an original pool of 21) summarized in table 1 that were discussed in Bloch (2003), De Serres and Groleau (1997), De Serres et al. (2003), Moschkovich (1992), Moschkovich et al. (1993), Rojano (2002), Schwarz and Yerushalmy (1992), Vergnaud, Cortes and Favre-Artique (1988) and Yerushalmy and Chazan (2002). For each item, a value has been assigned to each of the eight variables and the success rate has been computed. A standard multiple regression analysis has then been applied to the data with SPSS software. It is important to note that the available data did not permit modern item response theory analysis (such as Rash model). These analyses are expected to be done in the next steps of the project involving more items, more students and more complete data from international studies.

Table 1. Variables for the prediction of the difficulty of items related to linear functions

Name	Description	Possible values
PrsncY	Explicit presence of the y-intercept value	Absence, Presence
PrsncS	Explicit presence of slope parameter value	Absence, Presence
SgnY	Sign of the y-intercept value	Negative, Null, Positive
SgnS	Sign of the slope parameter value	Negative, Null, Positive
CmplxDR	Complexity of data representations	Easy, Moderate, Difficult
CmplxRS	Complexity of resolution skill	Concept, Application, Resolution
NbE	Number of equations	One, Two, Three
TypeA	Type of expected answer	Multiple choice, Short, Long

Results

The proposed model summarized in tables 2 and 3 was able to predict the success rate with a correlation coefficient *(r)* of 0.78 and an adjusted determination coefficient (r^2) of 0.50. These correlations are comparable to the upper limit of those presented in Hornke (2002) for high cognitive functions. The four variables that had the most important regression coefficients were *PrsncY* (the explicit presence of the y-intercept value), *SgnS* (the sign of the slope parameter value), *NbE* (the number of independent equations necessary to the solution) and *CmplxDR* (the complexity of data representations). Partial correlation coefficient for each variable is also presented in table 3. One of the particularities of the model is that it has been explicitly designed to be usable for classification of existing items but also for generation of new items.

Table 2. Linear determination coefficients for the proposed model

r	r^2	Adjusted r^2	Estimated standard error
0.78	0.61	0.50	0.14

Table 3. Linear regression coefficients for the proposed model

Name	Possible values	Correlation coefficient *(r)*	Regression coefficient *(b)*	Estimated standard error
(Constant)			0.15	0.22
PrsncY	0, 1	0.39	0.16	0.05
SignS	1, 2, 3	0.37	0.13	0.03
NbE	1, 2, 3	0.35	0.09	0.05
CmplxRS	1, 2, 3	-0.27	-0.01	0.05
TypeA	1, 2 ,3	-0.26	-0.05	0.04
CmplxDR	1, 2, 3	-0.19	-0.10	0.05
PresncS	0, 1	0.16	-0.06	0.05
SignY	1, 2 ,3	-0.01	0.03	0.05

As a proof of concept, the proposed model has then been used to generate about 3 000 nonequivalent new items with predicted success rates ranging from 0% (very difficult task) to 84% (very easy task). The generated items were constructed with the following structure: A) consider this situation, B) with these initial parameters, C) and this initial data representation, D) find this kind of answers, E) using this final data representation. An example of the generated items is presented in figure 1.

In Istanbul, a man is selling scarfs. This morning, he had a total of 200 scarfs. After 4 hours of constant selling, he had 150 scarfs left. What is the equation for the number (y) of scarfs left as a function of the time (x) in hours?

Figure 1. A generated item.

The item in figure 1 was generated with predictive variables set to the following values:

- PrsncY = 1 (y-intercept value is explicit);
- SignS = 1 (sign of the slope parameter is negative);
- NbE = 1 (one equation is necessary to get the solution);
- CmplxRS = 2 (the complexity of the resolution skill is moderate: application);
- TypeA = 2 (type of expected answer is short);
- CmplxDR = 2 (complexity of data representations is moderate);
- PresncS = 0 (slope parameter value is not explicit);
- SignY = 2 (sign of the y-intercept value is positive).

The expected success for this item rate can be computed from table 3:

$$0.15 + 0.16*1 + 0.13*1 + 0.09*1 + -0.01*2 + -0.05*2 + -0.10*2 + -0.06*0 + 0.03*2 = 27\%.$$

Conclusions and Implications

Predicting the difficulty of newly generated evaluation tasks could be an important part of a really useable implementation of computerized adaptive evaluation in science. This research has proposed a preliminary model with a coefficient of correlation of 0.78 that has been used to generate about 3 000 nonequivalent items of known difficulty related to linear functions for secondary students. This model could be used online by science teachers and researchers to generate new evaluation tasks or by computerized systems for adaptive evaluation. The next steps of this project will involve more items, more students, a broader subject in science and more complete data from international studies. As more predictive models for evaluations tasks will be proposed by research in this emerging field of item generation in science education, the authors plan to incorporate them in the online application developed for this project.

References

Bejar, I.I., Morley, M.E. and Benett, R.E. (2003). A feasibility study of-the-fly-item generation in adaptive testing. *The Journal of technology, learning and assessment*, 2(3), 1-29.

Bloch, I. (2003). Teaching functions in a graphic milieu: What forms of knowledge enable students to conjecture and prove? *Educational Studies in Mathematics*, 52, 3-28.

De Serres et Groleau (1997). *Mathématiques et langages*, Collège Jean de Brébeuf. Direction pédagogique, service de la recherche, 229 p.

De Serres, M., Bélanger, M., Piché, M. C., Riopel, M., Staub, C. et de Grandpré, C. (2003). *Intervenir sur les langages en mathématiques et en sciences*, Montréal : Modulo, 390 p.

Gierl , M. J., Zhou, J. and Alves, C. (2008). Developing a Taxonomy of Item Model Types to Promote Assessment Engineering *The Journal of Technology, Learning, and Assessment*, 7(2), 4-50.

Hively, W., Patterson, H. L. and Page, S. (1968), A "universe-defined" system of arithmetic achievement tests, *Journal of educational measurement*, no 5, p. 275-290.

Hornke, L. F. (2002), Item-generation models for higher cognitive function, in S. H. Irvine & P. C. Kyllonen (Eds), *Item generation for test development*, Mahwah, New Jersey, Lawrence Erlbaum Associates, p. 159-178.

Irvine, S. H. & Kyllonen, P. C. (2002), *Item generation for test development*, Mahwah, New Jersey, Lawrence Erlbaum Associates, 411 p.

Moschkovich, J. N. (1992). *Making Sense of Linear Equations and Graphs: An Analysisof Students' Conceptions and Language Use.* Dissertation, University of California at Berkeley.

Moschkovich, J., Schoenfeld, A. et Arcavi, A. (1993). Aspects of understanding: On multiple perspectives and representation of linear relations, and connections among them, in T.A. Romberg, E. Fennema and T.P. Carpenter (Eds.), *Integrating research on the graphical representation of function*. Erlbaum, Hillsdale, NJ, 69-100.

Riopel, M., Raîche, G., Pilote, M. and Potvin, P. (2009). La production automatisée de tâches d'évaluation. Dans Blais, J.-G. (Ed.) *Éducation et TIC en éducation/formation: enjeux, modèles et applications*. Ste-Foy, Québec: Presses de l'Université Laval, 309 p.

Rojano T. (2002). Mathematics learning in the junior secondary school: students' access to significant mathematical ideas. *Handbook of international research in mathematics education*, LEA, 806 p.

Schwarz, J. and Yerushalmy, M. (1992). "Getting students to function in and with algebra", in G. Harel and E. Dubinsky (Eds). *The concept of function: Aspects of epistemology and pedagogy* (MAA Notes). Mathematical Association of America, Washington, DC, Vol. 25, 261-289.

Vergnaud, G., Cortes A. et Favre-Artigue (1988). *Introduction à la didactique des sciences et des mathématiques. Problèmes épistémologiques et didactiques*. Presses Universitaires de France, 1993, 398 p.

Yerushalmy, M. et Chazan D. (2002). Algebra: curricular, graphing, and research. Flux in school algebra: Graphing technology, and research on student learning and teacher knowledge. *Handbook of International Research in Mathematics Education*, LEA, 806 p.

ASSESSMENT OF SCIENTIFIC INQUIRY SKILLS

Maik Walpuski
University of Osnabrück

Abstract

In this symposium, the testing of scientific inquiry competence with paper-and-pencil tests in large scale assessments was presented from different points of view. The evaluation of the German Educational Standards was explained using the example of scientific inquiry in comparison to the evaluation of Standards in Switzerland. While in Switzerland the model of competence was evaluated before standards for science were defined so data from the evaluation could be used for this purpose, the standards in Germany are normative standards which were formulated based on theory and will be assessed in 2012 after a pilot study in 2009. The underlying models of competence as well as first results from studies in both countries are presented in this paper.

Framework of the symposium

In response to findings of international large-scale assessments such as TIMSS and PISA the Assembly of German Ministers of Education agreed on educational standards as normative guidelines for secondary schools. In addition to *content knowledge*, *communication* and *evaluation & judgement*, *acquirement of knowledge* (e.g. through scientific inquiry) represents one area of competence of these standards for the three science subjects: biology, chemistry and physics. Tests for all areas of competence are currently being developed in the German project ESNaS (Evaluation der Standards in den Naturwissenschaften für die Sekundarstufe I – Evaluation of the National Educational Standards for Natural Sciences at the Lower Secondary Level). The items will be used in a nationwide pilot study in 2010 and in a nationwide assessment in 2012. At the same time, the Swiss project HarmoS (Harmonization of compulsory school) initiated by the Swiss Conference of Cantonal Ministers of Education (EDK: Erziehungs-direktorenkonferenz) has been completed after proposing to the policy makers performance standards for students at the end of grades 2, 6, and 9 in the four subjects mathematics, science, first language and second language based on the results from the evaluation. In this paper, we focus on the area of competence *acquirement of knowledge* (through scientific inquiry) for both countries. Tests from both projects measuring inquiry competences are discussed (ESNaS: Walpuski / Wellnitz, Hartmann & Mayer; HarmoS: Gut & Labudde). For the evaluation of HarmoS, amongst others an experimental test combined with a paper-and-pencil-test has been developed. Data from both tests will be presented. For the project ESNaS, data from a pilot study for the subject biology (Wellnitz, Hartmann & Mayer) will be presented with 46 items testing four defined inquiry skills *question*, *hypothesis*, *design* and *data*. In addition, data from related study (Mannel, Sumfleth & Walpuski) will be presented. Based on the "ESNaS model", a paper-and-pencil test has been designed especially for low achieving students. The items represent three inquiry skills: *generating a hypothesis* (fit between hypothesis and observation), *choosing an experiment* and *evaluating evidence*.

MODELLING SCIENTIFIC INQUIRY FOR LARGE SCALE ASSESSMENTS

Maik Walpuski
University of Osnabrück

Abstract

The German Educational Standards are divided into four areas of competence. All areas of competence are evaluated in a nationwide assessment. In the paper the assessment of the area of competence "acquirement of knowledge" is described. The test items are constructed using a model of competence which describes the "complexity" and the "cognitive processes" as aspects of tasks which present difficulty.

Introduction

As a result of large scale assessments such as TIMSS and PISA, Germany has shifted from input orientation to output orientation. As a consequence of the implementation of Educational Standards, tests for the evaluations of the standards have to be developed. Science educators from different universities are assigned to operationalise the German national standards in a model of competence which is the basis construction guidelines for the test items. The participating researchers are:

- For biology: Workgroup of Prof. Dr. Jürgen Mayer, University of Kassel
- For chemistry: Workgroup of Prof. Dr. Elke Sumfleth, University of Duisburg-Essen in cooperation with Prof. Dr. Maik Walpuski, University of Osnabrück
- For physics: Workgroup of Prof. Dr. Hans E. Fischer, university of Duisburg-Essen in cooperation with Prof. Dr. Alexander Kauertz, University of Education, Weingarten.

From the four defined areas of competence only the area of competence *acquirement of knowledge* is described in this paper. This area of competence again focuses on different subareas, which are *methods of scientific investigations* (scientific inquiry), *development of scientific models and theories* and *nature of science* as far as they are testable in paper-pencil-tests. In this paper different aspects of scientific inquiry (hypotheses, experiments, data) are focussed.

Rationale

As we know from literature, teaching and learning scientific inquiry in science education is not simple. Although experiments are an intrinsic part of science instruction, they are often used inadequately and structured by narrow directions resembling 'recipes' (Hofstein & Lunetta, 2004). As a result, "to many students, a 'lab' means manipulating equipment and not manipulating ideas" (Lunetta, 1998, p. 250). Even if students realize that experimental work is theory-based work, they often have deficiencies in doing experimental work. They often only test hypotheses for verification and do not try to falsify them (Lederman, 1992). Furthermore, students have difficulties in distinguishing between hypotheses, laws and theories (Lederman, 1992; Lederman et al., 2002) and they tend to accept scientific theories as the absolute truth (Lederman, 1992; Lederman et al., 2002; Tytler & Peterson, 2004). Deducing results from experiments presents another difficulty for students (Lederman et al., 2002). Students apparently do not learn scientific working and reasoning by simply practicing it (Bell, Blair, Crawford, &

Lederman, 2003), but have to be explicitly taught in the requested skills (Carey, Evans, Honda, Jay, & Unger, 1989; Schwartz, Lederman, & Crawford, 2004).

At the same time, German Educational Standards require students to be able to (e.g.)…

- identify and generate questions which can be answered with the help of scientific experiments.
- plan suitable experiments for testing hypotheses.
- carry out qualitative and quantitative analyses and to record them.
- find trends, structures and relations in data and to conclude from them. (KMK, 2005b).

The students' achievement according to the standards is to be evaluated in the ESNaS project (Evaluation der Standards in den Naturwissenschaften für die Sekundarstufe I – Evaluation of the National Educational Standards for Natural Sciences at the Lower Secondary Level). The ESNaS project has the following main goals:

- evaluation of the educational standards for the sciences (differentiated into the subjects biology, chemistry and physics)
- use of a shared model of competence for all three sciences for the construction of test items
- development of approx. 120 items per area of competence for a large scale assessment (paper-pencil-test)

For the use in a large scale assessment, the tests are limited to a paper-pencil-test which means that practical aspects of scientific inquiry cannot be investigated.

Methods

As already mentioned, the German Educational Standards are divided into four areas of competence: *content knowledge*, *acquirement of knowledge*, *communication* and *evaluation & judgement*. The area of competence *acquirement of knowledge* which this paper focuses on can be subdivided into three subareas which are *methods of scientific investigations* (scientific inquiry), *development of scientific models and theories* and *nature of science*. This paper focuses on the *methods of scientific investigations* only, which means dealing with scientific questions, hypotheses, experiments and data. Results from video studies show that a differentiating model of competence is required to analyse students' abilities. For example, students often do not formulate predictive hypothesis such as "If we change X, Y will change too" but formulate simple ideas like "We have to change X" (Walpuski, 2006), which is a kind of "low-level-hypothesis". For this reason, a model of competence, which originally was developed to describe physics content knowledge (Kauertz & Fischer, 2006), has been adapted for describing competences in science in general and combined with aspects of a model of inquiry competence (Mayer, Grube & Möller, 2008). The model of competence is the framework for the construction of test items in the ESNaS model.

The model (fig. 1) is a three-dimensional model with the dimensions content, complexity and cognitive processes. The dimension *areas of competence* distinguishes *use of content knowledge*, *acquirement of knowledge*, *communication* and *evaluation and judgement*. Complexity describes how much information has to be handled in a task. The general description shown in fig. 1 has to be differentiated for the different areas of competence. For the area of competence *acquirement of knowledge* and the subdimension *methods of scientific investigations* the complexity is defined as shown in tab. 1. The third dimension, cognitive processes, describes how students have to process information.

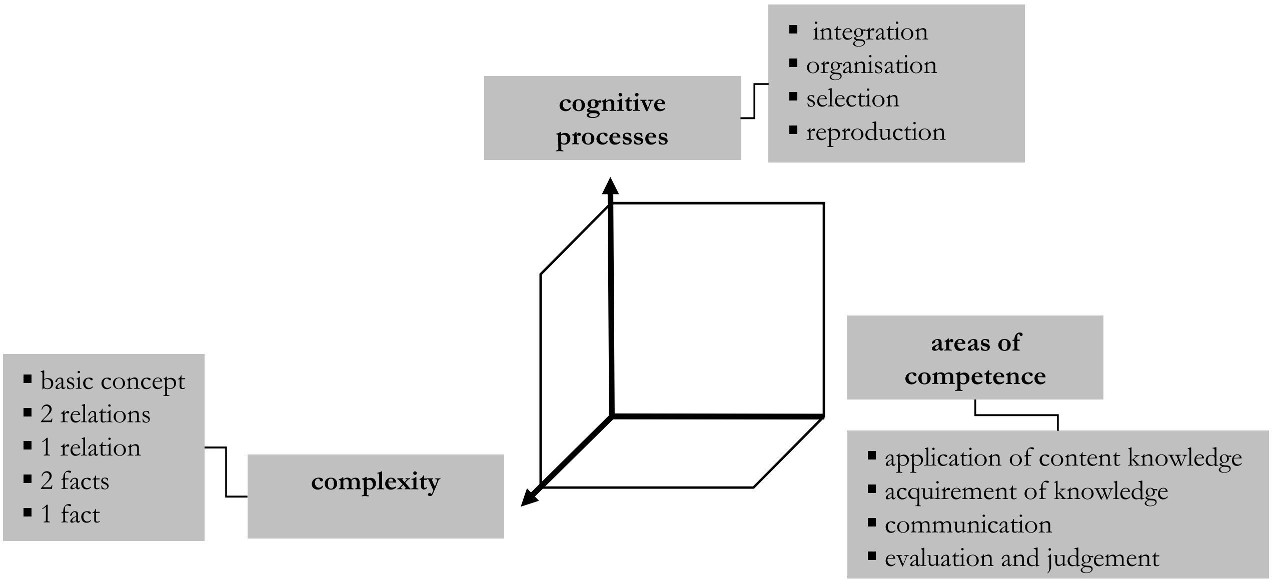


Fig. 1: Model of competence

Using this model, a group of teachers from all German states develops approx. 150 items of different formats (MC, short answer, extended response). These items are crosschecked by other teachers before they are reviewed by experts from different universities. The reviewed items are tested on psychometrical aspects in a pilot study before they are used for the evaluation of the standards.

Tab 1: Levels of complexity for *methods of scientific investigations*

complexity	definition
V basic (generic) concept	general characteristics of scientific investigations
IV 2 relations	interrelation between two variables, considering a control variable
III 1 relation	interrelation between two variables
II 2 facts	2 variables (no interdependency between variables)
I 1 fact	1 variable (dependent variable, independent variable, control variable)

Example and results

As shown by the example tasks (fig. 2) the situation in described in the item stem, because it is impossible to use hands-on experiments in a nationwide assessment. In the different items the students have to decide

- whether a hypothesis is testable or not
- whether a hypothesis fits a question or not
- whether an experiment fits a hypothesis or not
- which variables have to be assessed, controlled etc.
- how the data from an experiment can be interpreted.

Peter wants to find out if the reaction rate depends on the temperature (T) and/or the concentration (c) of the reactant. He already knows that the reaction rate is influenced by the degree of dispersion.

In an experiment he puts hydrochloric acid (c = 0,1 mol/L) into 4 test tubes. He places each test tube in a separate water quench, which have different temperatures (T = 10 °C, T = 20 °C, T = 40 °C, T = 80 °C). Afterwards, he adds zinc stripes of the same size to each of the test tubes and measures the time until the zinc stripes have completely dissolved.

Selection of one relation

Which question can Peter answer with his experiment? Mark with a cross.

Does the reaction rate depend on…

❑ the concentration of the acid?

❑ the temperature?

❑ the degree of dispersion of zinc?

Figure 2: Example task

Until now only first data from pre-pilot studies have been analysed, which means that there are no statistically reliable data until now. First analysis show the tendency that the items fit the model and that item difficulty can be estimated by the two dimensions *complexity* and *cognitive processes*.

Conclusions and Implications

In future, more reliable analyses will be possible with the data from the pilot study with approx. 10.000 students in 2010. Together with the data from the area of competence *use of knowledge* (see paper from Alexander Kauertz in this book of proceedings) it will be possible to describe students' competences for different areas of competence in a large scale assessment. Additional studies will have to be conducted in order to compare the results from large scale assessment with the performance of students in "real" inquiry situations and for further validation of the model.

References

Bell, R. L., Blair, L. M., Crawford, B. A. & Lederman, N. G. (2003). Just Do It? Impact of a Science Apprehenticeship Program on High School Students' Understandings of the Nature of Science and Scientific Inquiry. *Journal of Research in Science Teaching,* 40, 487-509

Carey, S., Evans, R., Honda, M., Jay, E. & Unger, C. (1989). 'An experiment is when you try it and see if it works': a study of grade 7 students' understanding of the construction of scientific knowledge. *International Journal of Science Education, 11,* 514-529.

Hofstein, A. & Lunetta, V. N. (2004). The Laboratory in Science Education: Foundations for the Twenty-First Century. *Science Education*, 88, 28-54.

KMK / Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2005). *Beschlüsse der Kultusministerkonferenz - Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss. Beschluss vom 16.12.2004*. [Decisions of the Assembly of German Ministers of Education - Educational standards in chemistry for lower secondary school. Decision of 16.12.2004] München, Neuwied: Luchterhand.

Lederman, N. G. (1992). Students' and Teachers' Conceptions of the Nature of Science: A review of the Research. *Journal of Research in Science Teaching, 29,* 331-359.

Lederman, N. G., Abd-El-Khalick, F., Bell, R. L. & Schwartz, R. S. (2002). Views of Nature of Science Questionnaire: Toward Valid and Meaningful Assessment of Learners' Conceptions of Nature of Science. *Journal of Research in Science Teaching, 39,* 497-521.

Lunetta, V. N. (1998). The school science laboratory: Historical perspectives and contexts for contemporary teaching. In: B. J. Fraser & K. G. Tobin (Eds.). *International Handbook of Science Education* (S. 249-262), Dordrecht: Kluwer.

Mayer, J., Grube, C., & Möller, A. (2008). Kompetenzmodell naturwissenschaftlicher Erkenntnisgewinnung. [A competence model of scientific inquiry] In U. Harms & A. Sandmann (Eds.), *Lehr- und Lehrnforschung in der Biologiedidaktik – Ausbildung und Professionalisierung von Lehrkräften* (pp. 63-79). Innsbruck: Studienverlag.

Schwartz, R. S., Lederman, N. G., & Crawford, B. A. (2004). Developing Views of Nature of Science in an Authentic Context: An Explicit Approach to Bridging the Gap Between Nature of Science and Scientific Inquiry. *Science Education*, 88, 610-645.

Tytler, R. & Peterson, S. (2004). From "Try It and See" to Strategic Exploration: Characterizing Young Children's Scientific Reasoning. *Journal of Research in Science Teaching*, 41, 94-118.

Walpuski, M. (2006). *Optimierung von experimenteller Kleingruppenarbeit durch Strukturierungshilfen und Feedback.* [Improving experimental collaborative working in small groups using structuring aids and feedback]. Berlin, Logos Verlag. Studien zum Physik- und Chemielernen. Niedderer, Hans; Fischler, Helmut & Sumfleth, Elke.

DEVELOPING A PAPER-AND-PENCIL-TEST TO ASSESS STUDENTS' SKILLS IN SCIENTIFIC INQUIRY

Nicole Wellnitz
University of Kassel

Stefan Hartmann
University of Kassel

Jürgen Mayer
University of Kassel

Abstract

German students have difficulties in the field of scientific problem solving and understanding scientific methods. To improve the quality of instructions for German students in biology, national educational standards were established in 2004. These standards will be evaluated on a regular basis through nationwide longitudinal assessments beginning in 2012. Scientific inquiry will be one area of the evaluation. Tests are currently being developed in the German project ESNaS (Evaluation of the National Educational Standards for Natural Sciences at the Lower Secondary Level). This study presents features from a pre-pilot study and investigates (1) whether item parameters are in accordance with hypothesized assumptions, (2) whether inquiry competence can be differentiated into the four skills „question", „hypothesis", „design" and „data" (as found in BIK 2008), and (3) whether this study confirms the lack of competence in scientific inquiry, which was found in PISA 2003 and 2006. Within the ESNaS-project, an instrument with 46 open and multiple-choice items was tested in a multi-matrix design (N = 293; 9th grade) and transferred into a four-dimensional Rasch-model. Results show that the four skills build a competence cluster "scientific inquiry". Students have less problems formulating scientific questions than formulating hypotheses, testing them and interpreting data. The test is slightly too difficult.

Introduction

In response to findings of international student assessments (TIMSS and PISA), the German government has agreed on educational standards as normative guidelines for secondary schools. Starting in 2012, these standards will be evaluated by means of nationwide assessments on a regular basis (KMK, 2006). Scientific inquiry represents one sub-area for biology (KMK, 2005).

To create an environment that encourages students to participate in science processes, it is important to find out how inquiry competence can be described and assessed. A theoretical framework and precise definition of competence are needed as a basis for valid performance measurement, which allows a reliable differentiation into sub-skills. According to international educational standards and curricula (DfES & QCA, 1999; NRC, 1996), scientific inquiry includes several relevant skills. However, concepts or the number of postulated skills differ (e.g., "posing questions", "generating hypotheses", "planning investigations", "identifying and controlling variables", "collecting data" or "data analysis"), which are mainly described as problem-solving (Gott & Duggan, 1998; Hammann et al., 2008; Klahr, 2000; Mayer, 2007).

Mayer, Grube and Möller (2008) found out that scientific inquiry can be divided into four central skills: “formulating questions”, “generating hypotheses”, “planning of investigation” and “interpreting data”. This study was associated with the German nationwide project BIK (Bayrhuber et al., 2007). In contrast to the presented study, Mayer et al. (2008) used a paper-and pencil-test with open-end test items in a sample of 1800 students (age 10 to 16), which were transferred into a four-dimensional partial-credit-model. The items represent four predicted skills, which all refer to the research method experimentation. The results show that a four-dimensional model for experimenting is more consistent than an one-dimensional model (Grube et al., 2007). Each skill is hierarchically divided into five competence levels (Möller et al., 2008).

These empirical findings provide the basis for the up-coming instrument to test students’ scientific inquiry competence by means of a large scale assessment. For this assessment the model of competence from Mayer, Grube and Möller (2008) and a model of competence from Kauertz and Fischer (2006) were adapted for describing competences in scientific inquiry. A multi-dimensional competence model**,** which was designed in the cooperation project ESNaS (Evaluation of the National Educational Standards for Natural Sciences at the Lower Secondary Level) (Walpuski et al., 2008) will be used to evaluate German educational standards. The model allows to describe and to analyse the structure as well as the levels of competence in specific domains.

Rationale

A test with 170 items for the sub-area *scientific inquiry* is currently being developed in the German project ESNaS. For this pre-pilot study just 46 items are of a special interest to answer the following questions:

1. Are item parameters (item discrimination, item difficulty etc.) and reliabilities in accordance with hypothesised assumptions?

2. Do the results of the actual study match the findings of other studies?
   a) Can inquiry competence empirically be divided into those four skills „question“, „hypothesis“, „design“ and „data“ as BIK 2008?
   b) Does this study evaluate the same lack of competences in scientific inquiry as PISA 2003 and 2006?

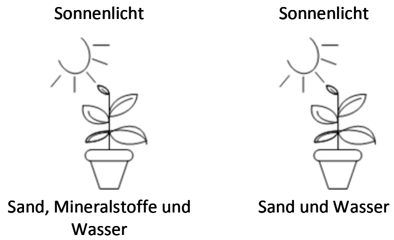
Design and methods

In order to develop a paper-and-pencil test to assess students’ competences in scientific inquiry, various inquiry methods are differentiated according to their specific hypothetical-deductive approach (Wellnitz & Mayer, 2008) and integrated in the ESNaS-competence model. The inner structure of observing, comparing and experimenting is adjusted to the necessary requirements. Consequently, each item is constructed specifically to test one of the four skills “question”, “hypothesis”, “design” and “data” (Figure 1). The four skills are divided into various levels of complexity and cognitive processes, each with separate difficulty degrees.

Based on the ESNaS-competence model a paper-and-pencil-test with multiple-choice and open test items was developed. The first items were tested in German schools with a sample of 293 students (9th grade) in a pre-pilot study. The paper-and-pencil test has the following characteristics: (1) among other items 46 multiple-choice and open test items were used to measure scientific inquiry; (2) each of those items represents one of the four skills: “question” (10 items), “hypothesis” (12 items), “design” (14 items) and “data” (10 items); (3) 30 minutes for the test items, another 10 minutes for descriptive data (e.g., age, gender, grade); (4) seven test booklets are elaborated in a multi matrix design.

Selecting of 1 fact to the aspect „hypothesis" (experimenting)

Your science teacher prepares an experiment explaining the growth of plants. She places two different objects for this purpose at the window:



Name one of the following hypotheses (assumptions) she want to test. Please check one of them!

The growth of a plant depends on …

□ …minerals.
□ …sunlight.
□ …sandy ground.
□ …water.

Figure 1. An example for a test item according to Baumert et al. (1998, p. 75); modified.

Test data is statistically analysed with methods of the item-response-theory (Rost, 2004). ConQuest (Wu et al., 1997) is used for statistic analysis to provide an indication of the compatibility of the model and the data. In order to test whether the four predicted skills can be discriminated and separated into four independent empirical dimensions, a one-dimensional Rasch-model as well as a four-dimensional model are compared regarding to their model fit.

Results

The analysis of item fit shows acceptable item-fit-data for 44 items ($.80 > wMNSQ < 1.20$; $T < 2.0$). Respectively, two items were not included in the test according to an unsatisfactory item fit.

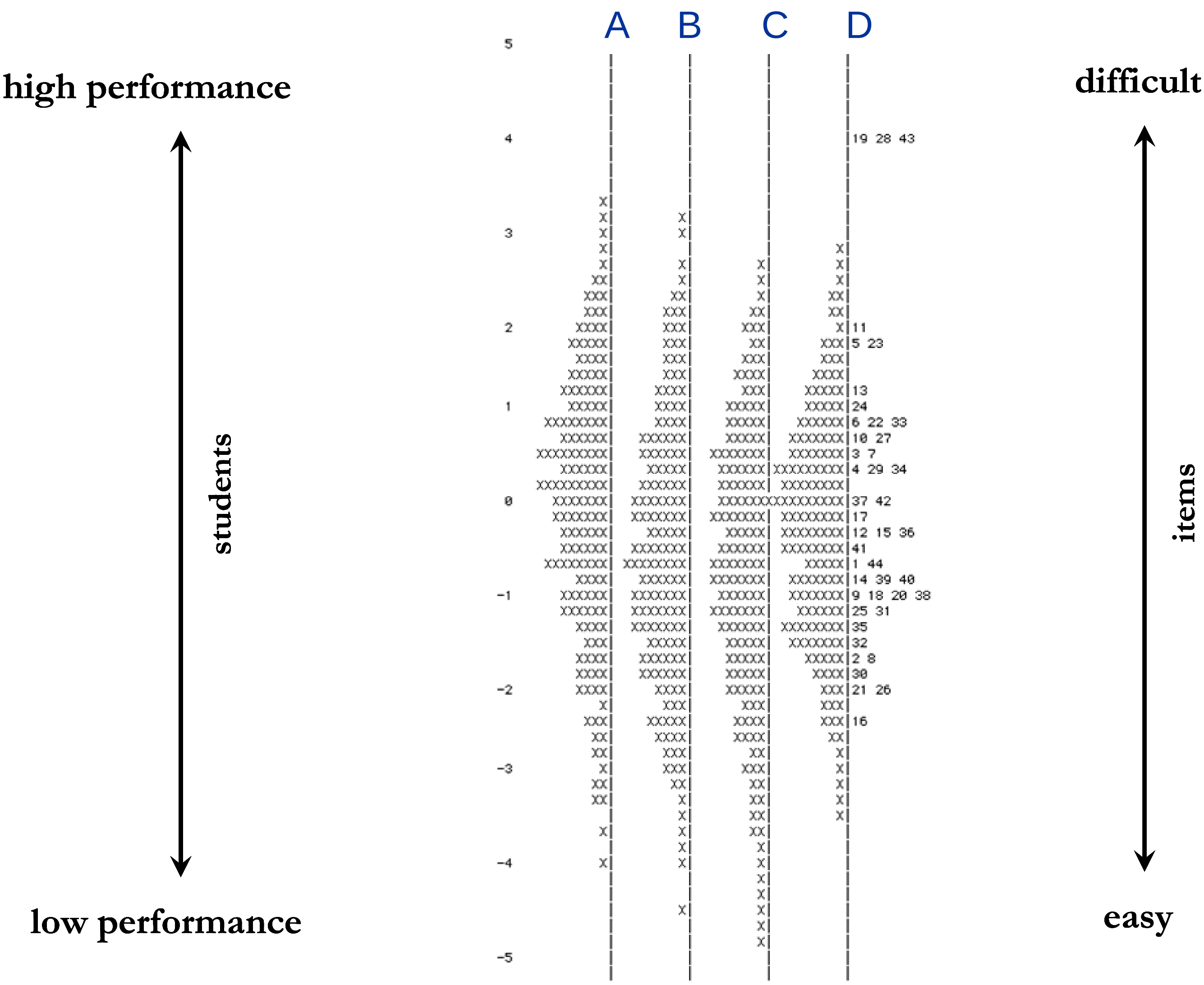
Comparison of model parameters indicates that the four-dimensional model is more consistent than the one-dimensional model. The differences between both models are significant (χ^2-test; $p<0.001$). These findings are supported by medium correlations found between the four skills (Table 1). Higher correlations would speak for a one-dimensional model, lower correlations for four independent skills.

Table 1. Inter-correlations and reliabilities

skills	(1)	(2)	(3)	(4)	Reliability
(1) question	–				.551
(2) hypothesis	.637	–			.545
(3) design	.690	.555	–		.538
(4) data	.425	.541	.341	–	.472

The reliabilities are low but generally acceptable (.472-.551). The consistency of the instrument requires a further improvement. One reason for the reliability could be the open-ended items with an average of one third of the whole test. Standardized tests using only multiple-choice items usually produce higher reliabilities (Stecher & Klein, 1997).

With the Rasch model it is possible to transform raw student scores and raw item difficulties to equal interval measures (Boone & Scantlebury, 2005). Figure 2 shows a Wright map; an aggregate map of all students' current proficiency levels versus all the item difficulties. Current proficiency levels and item difficulties are placed on the same logit scale. The most difficult items are on the top part of the scale, whereas easy items are located at the bottom. Students at the same level as an item have a 50% chance of correctly answering that item. Items above or below their ability level can also be answered correctly, but students have a more or less 50% chance of correctly answering the item (Boone & Scantlebury, 2005).



A: question, B: hypothesis, C: design, D: data

Figure 2. Person-item-map. Students are represented by the X sign. Each X sign represents 1.8 cases. Each of the test items is represented by the numbers 1 to 44.

The person-item map shows that there are some items (19, 28 and 43) which were way too difficult for all participating students. On the other side there were not enough items which were easy enough for everyone. Results show also that the four skills, „question", „hypothesis", „design" and „data" differ in variance and average difficulty. Items referring to "questions" are answered with a higher probability. This was not the case in the PISA study of 2006. The findings of PISA 2006 have shown that German students have problems in "identifying scientific issues" and there is also only a small percentage (9.0%) of students who are capable of carrying out the "identifying scientific issues" tasks at the two highest levels (OECD, 2006).

Conclusions and Implications

Results show that it was possible to develop reliable test items for the four scientific inquiry skills "question", "hypothesis"; "design" and "data". As in BIK 2008, which used a different method of measurement, the four different skills again could be differentiated. These findings confirm the assumption that scientific inquiry can be divided into these four skills. Correlations between the four dimensions indicate that those skills are based on one common competence, forming one cluster of competence.

It seems that a test instrument with open-end and multiple-choice items and the differentiation of the four skills could also work for different scientific methods like "experimenting", "observing" or "comparing". The named four skills differ in their level of difficulty. Students have less problems formulating questions than formulating hypotheses, testing them and interpreting data. Results show that the test is slightly too difficult.

The findings of four main inquiry skills for various scientific methods may provide the basis for a novel kind of individual assessment and the following intervention of students' inquiry skills in science education.

References

Baumert, J., Lehmann, R., Lehrke, M., Schmitz, B., Clausen, M., Hosenfeld, I., et al. (1998). *TIMSS - Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Deskriptive Befunde* [TIMSS – Mathematics and science education: an international comparison]. Opladen: Leske + Budrich.

Bayrhuber, H., Bögeholz, S., Eggert, S., Elster, D., Grube, C., Hössle, C., et al., (2007). Biologie im Kontext? Erste Forschungsergebnisse. [Biology in Context? First results]. *Der mathematische und naturwissenschaftliche Unterricht*, 60 (5), 304-313.

Boone, W. J., & Scantlebury, K. (2005). The Role of Rasch Analysis When Conducting Science Education Research Utilizing Multiple-Choice Tests. Science Education, 90, 253-269.

DfES / Department for Education and Skills, & QCA / Qualification and Curriculum Authority (1999*). Science. The National Curriculum for England*. London: HMSO.

Gott, R., & Duggan, S. (1998). Understanding Scientific Evidence – Why It Matters and How It Can Be Taught. In M. Ratcliffe (Ed.), *ASE Guide to Secondary Science Education* (pp. 92-99). Hatfield: Association for Science Education.

Grube, C., Möller, A., & Mayer, J. (2007). Dimensionierung wissenschaftsmethodischer Kompetenz. [Modelling inquiry competence] In H. Bayrhuber, F. X. Bogner, D. Graf, H. Gropengießer, M. Hammann, U. Harms et al. (Eds.), *Ausbildung und Professionalisierung von Lehrkräften. Internationale Tagung der Fachgruppe Biologiedidaktik im VBIO* (pp. 31-34). Kassel: IPN.

Hamann, M., Phan, T. T. H., Ehmer, M., & Grimm, T. (2008). Assessing pupils' skills in experimentation. *Journal of Biological Education*, 42 (2), 66-72.

Kauertz, A. & Fischer, H. E. (2006). Assessing Students' Level of Knowledge and Analysing the Reasons for Learning Difficulties in Physics by Rasch Analysis. In X.Liu & W. J. Boone (Eds.), *Applications of Rasch Measurement in Science Education* (pp. 212-246). Maple Grove, USA: JAM press.

Klahr, D. (2000). *Exploring science: The cognition and development of discovery processes.* Cambridge, MA: MIT Press.

KMK / Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2005). *Beschlüsse der Kultusministerkonferenz - Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss. Beschluss vom 16.12.2004*. [Decisions of the Assembly of German Ministers of Education - Educational standards in biology for lower secondary school. Decision of 16.12.2004] München, Neuwied: Luchterhand.

KMK / Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2006). *Gesamtstrategie der Kultusministerkonferenz zum Bildungsmonitoring.* [Strategy of the Assembly of German Ministers of Education for education monitoring] Bonn: LinkLuchterhand.

Mayer, J. (2007). Erkenntnisgewinnung als wissenschaftliches Problemlösen [Inquiry as scientific problem solving}. In D. Krüger & H. Vogt (Eds.), *Handbuch der Theorien in der biologiedidaktischen Forschung* (pp. 178-186). Berlin: Springer.

Mayer, J., Grube, C., & Möller, A. (2008). Kompetenzmodell naturwissenschaftlicher Erkenntnisgewinnung. [A competence model of scientific inquiry] In U. Harms & A. Sandmann (Eds.), *Lehr- und Lernforschung in der Biologiedidaktik – Ausbildung und Professionalisierung von Lehrkräften* (pp. 63-79). Innsbruck: Studienverlag.

Möller, A., Grube, C., & Mayer, J. (2008). Skills and Levels of Students' Inquiry Competence in Lower Secondary Biology Education (Grade 5-10). *Paper presented at the conference of NARST*. Baltimore, USA, April 2008: Paper published on the conference CD.

NRC / National Research Council (1996). *National science education standards*. Washington, DC: National Academy Press.

OECD (2006). *PISA 2006. Science Competencies for Tomorrow's World. Volume 1: Analysis*. In www. oecd.org/dataoecd/30/17/39703267.pdf (10.2. 2009).

Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion* (2. Aufl.). [Textbook test theory – Test construction] Bern: Hans Huber.

Stecher, B. M., & Klein, S. P. (1997). The cost of science performance assessments in large-scale testing programs. *Educational Evaluation and Policy Analysis*, 11 (1), 1-14.

Walpuski, M., Kampa, N., Kauertz, A., & Wellnitz, N. (2008). Evaluation der Bildungsstandards in den Naturwissenschaften. [Evaluation of educational standards in science]. *Der mathematische und naturwissenschaftliche Unterricht,* 61 (6), 323-326.

Wellnitz, N., & Mayer, J. (2008). Evaluation von Kompetenzstruktur und -niveaus zum Beobachten, Vergleichen, Ordnen und Experimentieren.[Evaluation of structure and levels of competence in observing, comparing, arranging and experementing] In D. Krüger, A. Upmeier zu Belzen, T. Riemeier, K. Niebert (Eds.), *Erkenntnisweg Biologiedidaktik 7* (pp. 129-143). Hannover:Universität Kassel.

Wu, M. L., Adams, R. J., & Wilson, M. (1997). *ConQuest - Generalised item response modelling software* (Draft Release 2). Camberwell: Australian Council for Educational Research.

ASSESSMENT OF STUDENTS' PRACTICAL PERFORMANCE IN SCIENCE: THE SWISS HARMOS PROJECT

Christoph Gut
University of Applied Sciences Northwestern Switzerland

Peter Labudde
University of Applied Sciences Northwestern Switzerland

Abstract

With the project HarmoS (Harmonization of compulsory school), the Swiss Conference of Cantonal Ministers of Education intends to harmonize and monitor the cantonal school systems in Switzerland. In October 2008, the project has been finished by proposing to the policy makers performance standards for students at the end of grades 2, 6, and 9 in the four subjects mathematics, sciences, first language and second language. In sciences, the proposed basic standards are based on a three-dimensional competence model that has been evaluated in a nationwide assessment, including a paper-and-pencil-test in 2007 with about 8000 students and an experimental test in 2008 with about 2000 students of the 2nd, 6th, and 9th year level. The results of the experimental tests show, that the correlation between the practical skill correlates highly (correlation coefficients higher than 0.70) with other skills.

Introduction

In October 2008, the Swiss project HarmoS (Harmonization of compulsory school) initiated by the Swiss Conference of Cantonal Ministers of Education (EDK: *Erziehungsdirektorenkonferenz*) has been finished by proposing to the policy makers performance standards for students at the end of grades 2, 6, and 9 in the four subjects mathematics, science, first language and second language. The proposed standards are handled as minimal or so-called basic standards and are based on competences as defined by Weinert (2001). With the project HarmoS, the EDK intends to harmonize and monitor the educational landscape in Switzerland (EDK, 2004a, 2004b), which is divided into 26 cantons, each having its own educational system until now. Other countries face similar problems and have started to develop and establish standards in science education (Waddington et al., 2007).

Rationale

In science, the standards are based on a competence model that includes three dimensions: skills, domains of contents, and achievement levels (Fig. 1). The model distinguishes eight skills addressing cognitive and social competences as well as skills for practical laboratory work, namely: “To show interest and to be curious”, “To ask questions and investigate”, "To exploit information sources”, “To organize, structure, and model”, “To assess and judge”, “To develop and implement”, “To communicate and exchange views”, and “To work self-reliantly and to reflect”. The domains of contents are within physical or living systems and are constructed like a STS-approach. For each skill, *a priori* achievement levels have been formulated, ranging consecutively from the 2nd to the 9th grade (Labudde, 2007). The competence model has been evaluated in a nationwide assessment, including a paper-and-pencil-test in 2007 with about 8000 students and an experimental test (lab-work and hands-on-activities) in 2008 with about 2000 students of the 2nd, 6th, and 9th year level. On the basis of these results and the evaluation of the competence model, standards in science have been formulated (HarmoS, 2008). In our paper we focus on the experimental tests and the validation of the competence model with respect to the experimental skill "To ask questions and investigate".

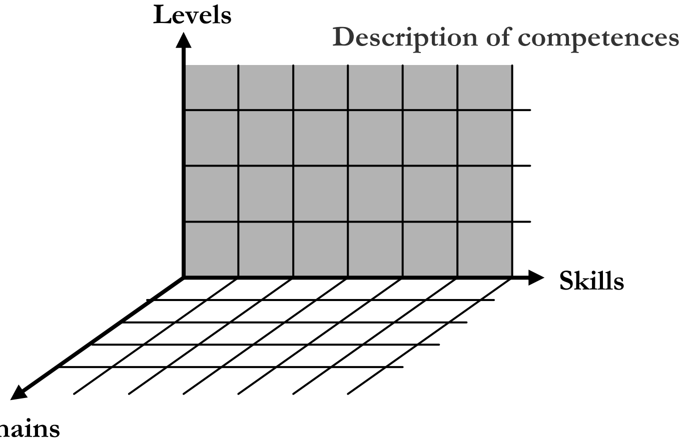


Figure 1. HarmoS competence model for sciences. The skills and the levels of the skills are described independently from the domains.

Test design

An essential feature of the HarmoS science project is the assessment of students' performance by large-scale experimental tests, according to the tradition from APU (1988), TIMSS (1997), and NAEP (2009). In spring 2008, an experimental test combined with a paper-and-pencil-test with 1468 students of the 6th and the 9th grade in the German, French and Italian speaking part of Switzerland was accomplished. Altogether 15 experimental and 35 paper-and-pencil tasks on topics from the domains "Planet earth", "Motion, force, energy", "Structure and properties of matter", "Plants and animals", "Ecosystems" and "Nature, society, technology: perspectives" have been used, including more than 300 items related to the skills „To ask questions and investigate", „To exploit information sources", „To organize, structure and model", „To assess and judge", and „To develop and implement". Almost half the items concern one of the five sub-dimensions of the skill "To ask questions and investigate": 1) "To observe consciously", 2) "To formulate questions, problems and hypotheses", 3) "To select and use tools, instruments and materials", 4) "To conduct investigations and experiments" and 5) "To reflect results, methods and techniques". Each student worked on two experimental und five paper-and-pencil tasks during two sessions of 90 minutes. During the tests the students were assisted with the handling of difficult devices, for instance with handling a microscope. In grade 6, in order to facilitate the understanding, the supervisors read the tasks loudly, whereas in grade 9 no such help was given.

Method and results

The test sheets were scored in June 2008 The statistical analysis followed in July and August 2008. After a first one-dimensional Rasch analysis of all experimental tasks and paper-and-pencil-tasks, about a third of the items was left for further multi-dimensional analyses due to insufficient item fits and due to large differential items functioning with respect to the various test languages.

In a second two-dimensional analysis the items of the experimental test (EXP) are compared with the items of the paper-and-pencil-test (PP). As not all items of the experimental tasks address the skill "To ask questions and investigate", further analysis is necessary. In a three-dimensional analysis the items of the experimental test (EXP) are split into items that address the experimental skill "To ask questions and investigate" and into items that address other skills. The experimental items are compared with the items of the paper-and-pencil test as well as with the non-experimental items of the experimental test. The results, i.e. the correlations coefficients, are shown in tables 1 and 2.

Generally, the various skills correlate highly at 0.70-0.79. In grade 6, the correlation between the experimental items and the non-experimental items, 0.70 and 0.71, is remarkably lower than the correlation between the non-experimental items. Unfortunately, no such clear effect can be measured in grade 9.

Interpretation

The practical problems and the scores of the students can be used to illustrate the skill "To ask questions and investigate" and the corresponding basic standards. They can serve as a useful tool for the implementation of standards.

The high correlations between the different skills lead to several questions: How far differ the skills in sciences from each other? Can a skill like "To ask questions and investigate" be really distinguished from other non-practical skills? Were the tests, their design and the practical problems suitable to measure practical (sub-)skills like "To observe consciously" or "To select and use tools, instruments and materials"? How far can a competence model be validated by tests like ours?

Table 1. Comparison of the 2- und 3-dimensional analyses for grade 6

Grade 6 (deviance = 36856)		
	PP	**EXP**
PP	Rel = .75 M = –.15 Var = .35	**0.77**
EXP		Rel = .75 M = –.11 Var = .37

Grade 6 (deviance = 36778)			
	PP	**EXP** other skills	**EXP** “To ask questions and investigate”
PP	Rel = .73 M = –.15 Var = .37	**0.79**	**0.71**
EXP other skills		Rel = .66 M = –.13 Var = .39	**0.70**
EXP “To ask questions and investigate”			Rel = .66 M = –.10 Var = .48

Table 2. Comparison of the 2- und 3-dimensional analyses for grade 9

Grade 9 (deviance = 36179)		
	PP	**EXP**
PP	Rel = .80 M = –.04 Var = .52	**0.77**
EXP		Rel = .79 M = –.14 Var = .67

Grade 9 (deviance = 36131)			
	PP	**EXP** other skills	**EXP** “To ask questions and investigate”
PP	Rel = .81 M = –.04 Var = .51	**0.75**	**0.74**
EXP other skills		Rel = .66 M = –.25 Var = .93	**0.72**
EXP “To ask questions and investigate”			Rel = .71 M = –.12 Var = .73

References

APU (1988). *Assessment of Performance Unit. Science at Age 15. A Review of APU Survey Findings 1980-1984*. London: HMS Office.

EDK (2004a). *HarmoS – Zielsetzungen und Konzeption*. Bern: EDK. Retrieved December 22, 2008, from http://www.edk.ch/d/EDK/Geschaefte/framesets/mainAktivit_d.html.

EDK (2004b), *Projekt HarmoS – Offertenausschreibung für Entwicklung von Konpetenzmodellen,* Bern: EDK. Retrieved Decembre 22, 2008. from http://www.edk.ch/d/EDK/Geschaefte/framesets/mainAktivit_d.html.

HarmoS (2008). *HarmoS Naturwissenschaften+: Kompetenzmodell und Vorschläge für Bildungsstandards. Wissenschaftlicher Schlussbericht*. Bern: EDK.

Labudde, P. (2007). How to Develop, Implement and Assess Standards in Science Education? 12 Challenges from a Swiss Perspective. In: D. Waddington, et al. (Eds.), *Making it comparable – Standards in Science Education* (pp. 277-302). Münster: Waxmann.

NAEP (2009). S*cience Assessment and Item Specifications for the 2009 National Assessment of Educational Progress*. Washington: National Centre for Education Statistics.

TIMSS (1997). *Performance Assessment in IEA's TIMSS*. Chestnut Hill: Boston College

Waddington, D., et al. (Eds.) (2007). *Making it comparable – Standards in Science Education*. Münster: Waxmann.

Weinert, F. H. (2001). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In: F. H. Weinert (Ed.), *Leistungsmessung in Schulen* (pp. 17-31). Weinheim: Beltz.

STUDENT ASSESSMENT IN THE AREA OF ACQUIREMENT OF KNOWLEDGE

Susanne Mannel
Univeristy of Duisburg-Essen

Elke Sumfleth
University of Duisburg-Essen

Maik Walpuski
University of Osnabrueck

Abstract

A *new curriculum in primary science education for North-Rhine-Westphalia (Germany) has come into effect with the onset of school year 2008 / 2009. The new curriculum defines aspects of scientific inquiry competence students should have attained by the end of primary school. According to the curriculum these aspects should be among the entry competencies of German 5th graders (10 – 11 year old).*

The project's main goal is to assess students' scientific inquiry competence at the beginning of secondary school (grade 5). Therefore, a paper-and-pencil test was developed for measuring students' achievement, whereas the project's main focus was on developing tasks for measuring low-performers competence. The test was conducted with students from different types of German secondary schools. For item construction a three-dimensional model of competence consisting of the dimensions of complexity, cognitive processes and previous knowledge was used. Moreover, the SDDS-model by Klahr & Dunbar (Klahr, 2000) was used for the development of different types of experimental tasks. The main part of data analysis was conducted by using Rasch Analysis.

Background, Framework, and Purpose

With the onset of school year 2008 / 2009 a new curriculum in primary science education for North-Rhine-Westphalia came into effect (Ministerium für Schule und Weiterbildung des Landes Nordrhein Westfalen, 2008). All in all four different areas of competence are defined by the new curriculum. One of them is the area of *acquirement of knowledge*. Aspects of scientific inquiry competence that should be mastered by the end of primary school are described within this area of competence.

Such a new curriculum needs to be evaluated. Therefore, one main goal of this project was the development of a student achievement test (paper-and-pencil test). Since the test should be adequate to measure differences in performance between students from different types of German secondary schools (*Hauptschule* (comparable to lower-secondary modern school) and *Gymnasium* (comparable to grammar school)), items are needed belonging to different performance or difficulty levels. According to this requirement, a model of competence, defining different performance levels was used for the development of test items. The model, used in this project is an adapted version of the ESNaS model (Walpuski, Kampa, Kauertz & Wellnitz, 2008). Three dimensions, namely complexity, cognitive processes and previous knowledge were taken into account (figure 1). Each dimension is stepped hierarchically, whereas the dimension complexity is built up by three factor levels and the dimensions cognitive processes and previous knowledge are built up by two factor levels (figure 2). An increase in item difficulty from the first to the second (to the third) factor level within the same dimension is expected to be found by the hierarchical structure of the model.

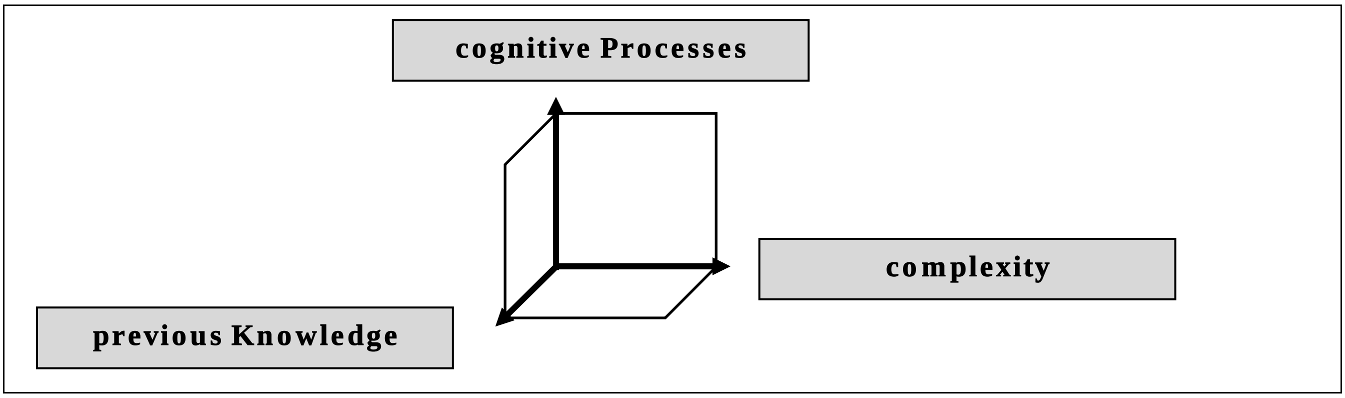


Figure 1. Adapted model of competency's dimensions.

Each area of the adapted model of competence is built up by exact one factor level of each of the model's three dimensions. Figure 2 shows the different areas of the adapted model of competence that are taken into account in this project. Each area is indicated by one cube exactly (figure 2).

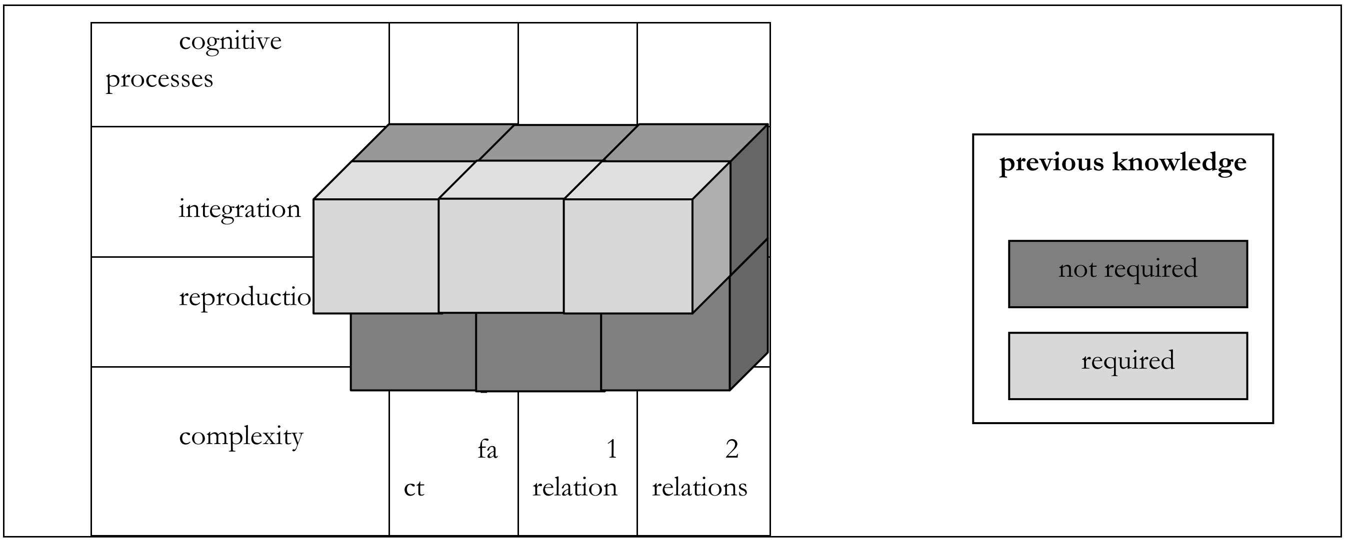


Figure 2. Areas of the adapted model of competency.

All in all 140 items were developed that are equally distributed into all different areas of the adapted model of competence, whereas a minimum of 12 items were developed for each area defined by the model.

According to the SDDS-model by Klahr & Dunbar (Klahr, 2000) and e.g. studies by Walpuski (2006), the focused *area acquirement of knowledge* is regarded with respect to three characteristic steps: *generating a hypothesis* (fit between hypothesis and observation), *choosing an experiment* and *evaluating evidence*. Therefore, each of the tasks is related to one of those three specific steps described by the SDDS-model. The following example is related to the type of task (characteristic step), namely *choosing an experiment* (figure 3). The area of the model the tasks is located in, is built up by the factor levels *1 relation* (dimension of complexity), *integration* (dimension of cognitive processes) and *previous knowledge required* (dimension of previous knowledge).

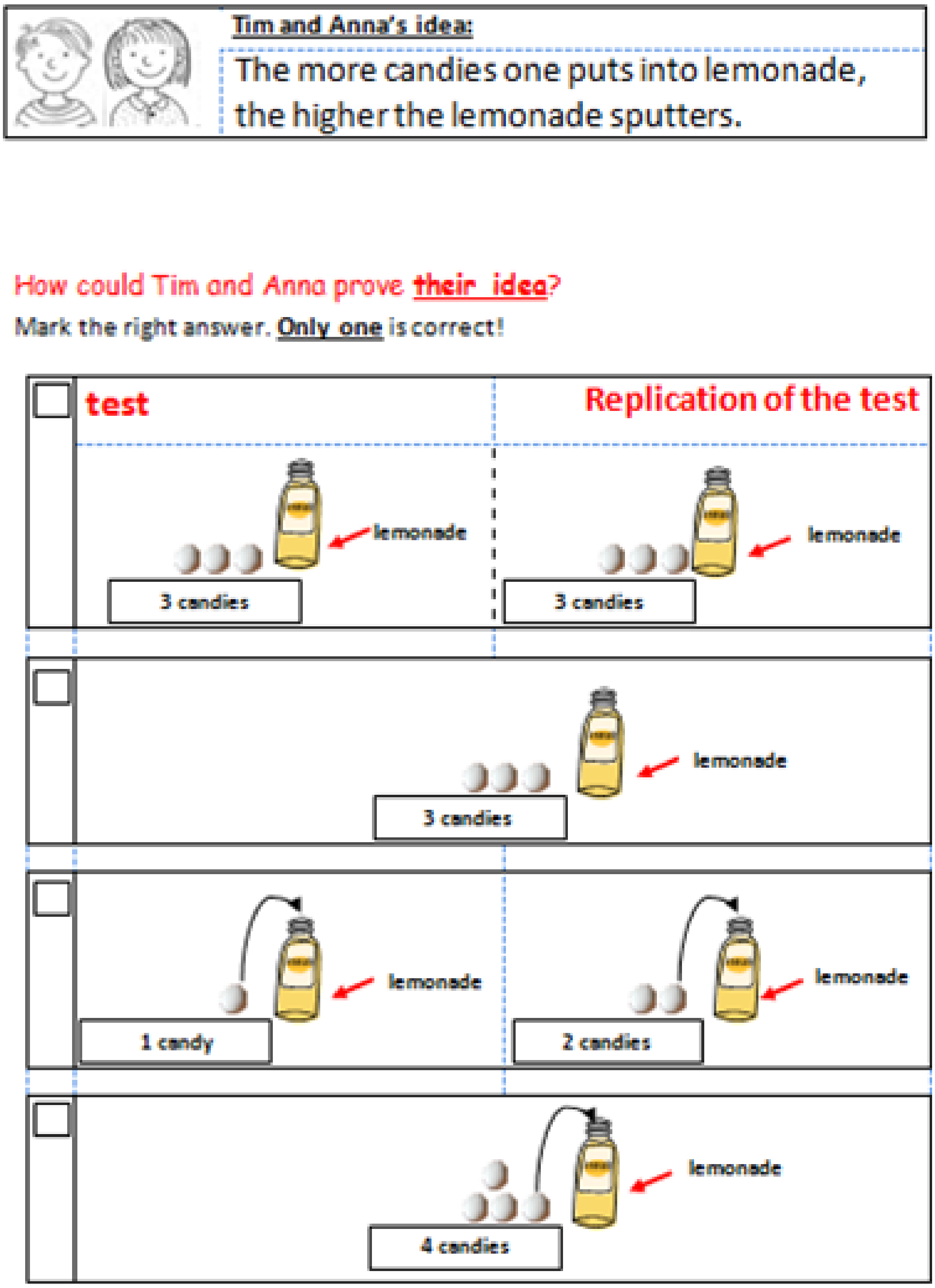


Figure 3. Task example (choosing an experiment).

Rationale

In his dissertation on difficulty-generating aspects, Kauertz (2008) showed a relation between the factor task complexity and item difficulty (p= 0,36, p < 0,001). Some indicators of a relation between task-related cognitive requirement (a cognitive process) and its difficulty are found in the work of Bloom (Bloom et al. 1956), Kottke and Schuster (1999) and Schabram (2007). According to these findings, the greatest influence on item difficulty was expected to be related to the factor of complexity. From here the first research question arose:

1. Can the proposed difficulty-generating aspects in tasks on experimental skills (scientific inquiry competence) be confirmed empirically?

According to the results of PISA 2003 (PISA-Konsortium Deutschland, 2004) differences in students' achievement which result by the type of school are expected to be found, too. Due to those results, the second research question arose:

2. Analogous to PISA, can significant differences between students from *Gymnasium* and *Hauptschule* be detected? If so, in which areas can these differences be found?

Methods

In order to prove whether or not the model's different factors have an impact on item difficulty in the predicted way, an item pool was developed that contained all in all 140 multiple-choice items that are distributed into the different areas defined by the adapted model of competence.

In the main study, conducted between August and November 2008, the item pool was tested by an entire sample of 1134 5th grade students consisting of students from the following two types of North-Rhine-Westphalia's schools: grammar school (Gymnasium) and lower secondary modern school (Hauptschule). Due to the number of tasks a multi-matrix-design was used to spread the items over 18 different test booklets, 18 items each, whereas every item is presented in a minimum of four different test booklets. An overlapping of items resulting from this approach allows IRT method for task analysis (Bond & Fox, 2001), therefore a comparison between item and person parameters is possible. Due to the amount of text that has to be read to work over the experimental tasks of the students' achievement test, a speed-test on reading comprehension (Lenhard & Schneider, 2006) was conducted to measure reading comprehension's influence on students' performance in the achievement test. Moreover, subscales of a test on cognitive skills (Heller & Perleth, 2000) were conducted to control the influence of cognitive skills on students test performance as well. In addition to that a test on students' content knowledge relating to primary school science was presented to a smaller part of the tested participants. Data collected by these additional tests were used for the process of validating the test on students' experimental skills.

Results

Results, presented here, are related to the main study. Currently, only data of the test on students' experimental skills are available. Data from all additional tests will be available by the end of November 2009. Firstly, a Rasch analysis was conducted for the total sample. Next the data was separately analyzed for both partial samples: "Gymnasiasten" (grammar school) and "Hauptschüler" (lower secondary modern school). The Rasch model could be confirmed concerning the total sample and the partial sample "Gymnasiasten". Concerning the total sample 121 of 140 items show a good fit to the Rasch model (.80 > MNSQ < 1.20; -2.00 > T < 1.80). Similar values are observed for "Gymnasiasten". With regard to the total sample, the estimators for the item difficulty are located between $\sigma = -3.16$ and $\sigma = 1.69$ (the mean is fixed at $\sigma = 0$). Concerning the partial sample of *Gymnasiasten*, these parameters range from $\sigma = -3.25$ to $\sigma = 2.09$. The mean person parameter is situated at -0.27 (total sample), 0.33 respectively (partial sample of *Gymnasiasten*). The reliability (EAP/PV reliability: .75 - .79) as well as the variance (1.032 – 1.279) are satisfactory. All in all the data indicate the test being slightly too difficult with regard to the total sample and slightly too easy concerning "Gymnasiasten". Differing from that are the results for "Hauptschüler". Here, item parameter, fit-parameter as well as the t-values initially indicate a good fit to the Rasch model while the reliability is poor (.29). In addition to that the difference between mean person ability (-0.86) and the estimated item difficulty ranges from $\sigma = -3.45$ to $\sigma = 1.63$ (the mean is fixed at $\sigma = 0$) and show clearly that the test is too difficult for this sample (figure 4). Merely 53 of overall 140 items are located within the spectrum of the participants' abilities (figure 4). Most of these items belong to the lowest level of complexity.

Nevertheless the findings show that the test instrument is adequate to measure differences between students performance within the partial sample of German Gymnasiasten and between students from both different types of school (Hauptschule and Gymnasium).To investigate the influence of all three factors (complexity, cognitive processes, procedural knowledge) on item difficulty correlations were calculated between complexity and item parameter (R^2 = .50, p < 0.001), between cognitive processes and item parameter (R^2 = .33, p < 0.001) as well as between procedural knowledge and item parameter (R^2 = .25, p < 0.001). As expected, all factors influence the difficulty of an item with task complexity showing the greatest influence.

To answer the second research question (Analogous to PISA, can significant differences between students from *Gymnasien* and *Hauptschulen* be detected?) correlations were calculated between the mean person ability between the two different types of schools. The difference between the mean person ability of students from Gymnasium (0.469) and mean person ability of students from Hauptschule (-0.915) is highly significant, whereas, as expected, most of the students from Gymnasium outperformed the students from Hauptschule (figure 5).

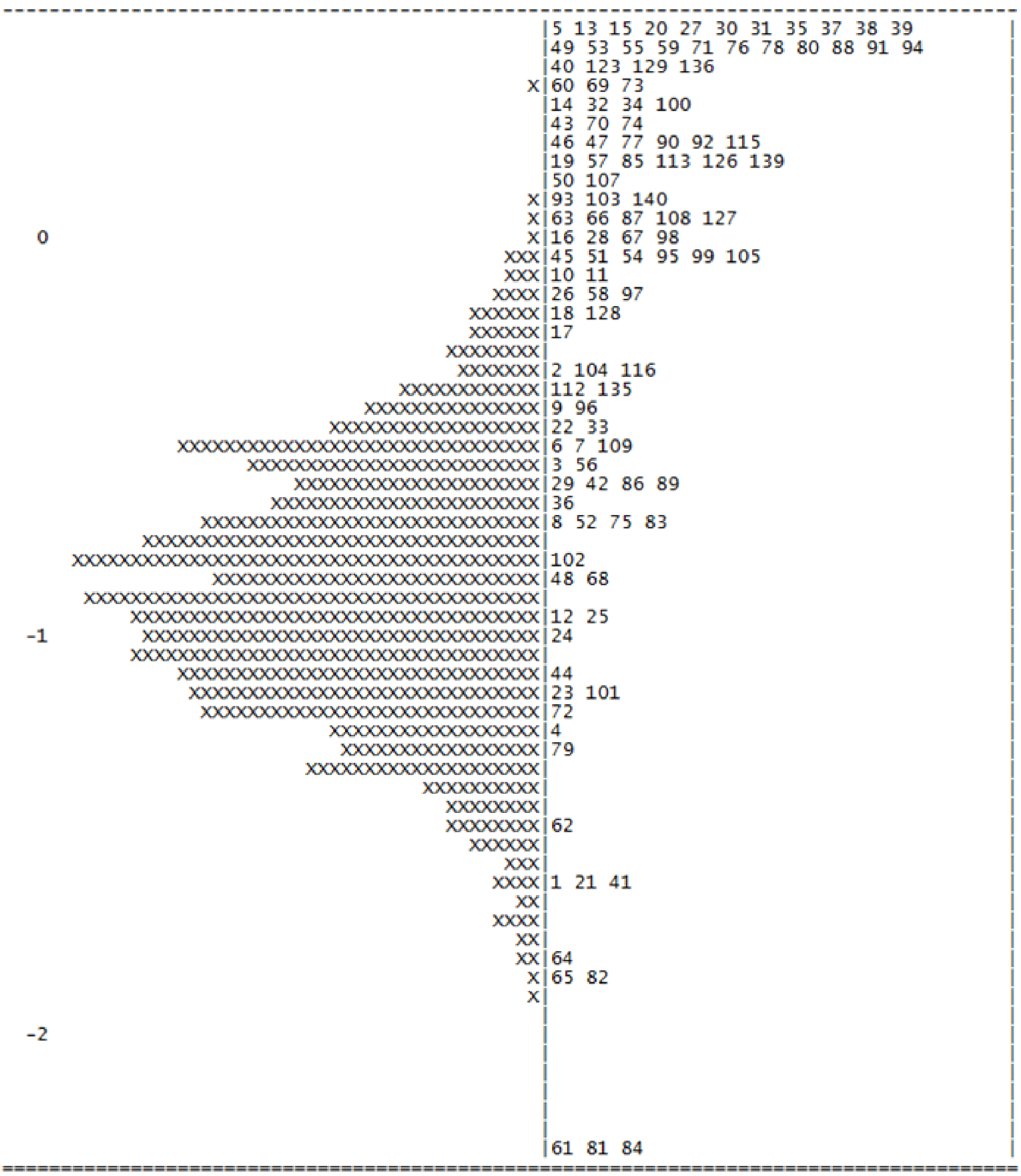


Figure 4. Graphical output (Wright map) related to the partial sample of German Hauptschüler (Rasch Analysis).

Conclusions and Implications

As shown by Rasch analysis the test on students' experimental skills is applicable for German Gymnasiasten. Moreover it allows differentiation between Gymnasiasten and Hauptschüler.

Furthermore, the main study's first results indicate that a test consisting of many more items located at the lowest level of complexity (1 fact) should be more adequate for measuring differences in the performance of German "Hauptschüler". The development and the evaluation of a test that met those requirements was the main goal of the follow-up study which occurred in spring 2009. Therefore, selective tasks of the main study that fitted into the spectrum of low-achievers ability were used as anchor items. Tasks that were not located within this spectrum were replaced by newly developed items; all of them on the complexity level of one fact. To ensure the comparability of both studies, main and follow-up, the number of items per test booklet as well as the test instruction and the conduction of the test were identical. The follow-up study took place in 10 classes of grade 5 (only German "Hauptschule").

	lower-secondary modern school (Hauptschule)	**grammar school (Gymnasium)**
mean person ability	-0.915	0.469

*** p < .001

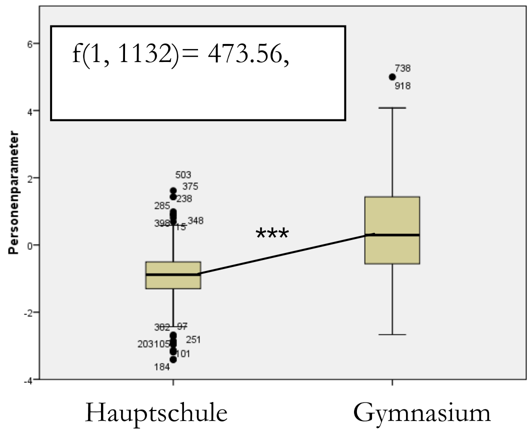


Figure 5. Correlation between mean person of ability of students from different types of school.

References

Bloom, B., Englehart, M. Furst, E., Hill, W., & Krathwohl, D. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain.* New York, Toronto: Longmans, Green.

Heller, K.A. & Perleth, C. (2000). Kognitiver Fähigkeitest für 4. Bis 12. Klassen, Revision. [Test on cognitive skills (grade 4 – 12)]. Göttingen: Beltz-Test.

Kauertz, A. (2008). Schwierigkeitserzeugende Merkmale physikalischer Leistungstestaufgaben [Difficulty-generating aspects of performance assessment tasks in physics]. Berlin: Logos Verlag.

Klahr, D. (2000). Exploring science: The cognition and development of discovery processes. Cambridge, MA: MIT Press.

Kottke, J. L., & Schuster, D. H. (1990). Developing tests for measuring Bloom's learning outcomes. *Psychological Report*, 66, 27-32.

Lenhard, W., & Schneider, W. (2006): Elfe 1-6 – Ein Leseverständnistest für Erst- bis Sechstklässler [Test on students' reading comprehension (grade 1-6)]. Göttingen: Hogrefe.

Ministerium für Schule und Weiterbildung des Landes Nordrhein-Westfalen (2008). Richtlinien und Lehrpläne für die Grundschule in Nordrhein-Westfalen. Deutsch, Sachunterricht, Mathematik, Englisch, Musik, Kunst, Sport, Evangelische Religion, Katholische Religion. [Guidelines and Curriculum for primary school in Northrhine-Westphalia for German, Science, Mathematics, English, Music, Arts and Religion], Retrieved October 1, 2008, from: http:// ritterbach.de/lp_online/2011.pdf

Schabram, K. (2006). Lernaufgaben im Unterricht: Instruktionspsychologische Analysen am Beispiel der Physik. [Learning tasks within lessons: Instructional psychological analysis in physics] (Doctoral dissertation, University of Duisburg-Essen, Northrhine-Westphalia, Germany2006). Retrieved from 08, January, from http://duepublico.uni-duisburg-essen.de/servlets/ DerivateServlet/Derivate-19040/Diss_Schabram.pdf

PISA-Konsortium Deutschland (Eds.). (2004). *PISA 2003: Der Bildungsstand der Jugendlichen in Deutschland - Ergebnisse des zweiten internationalen Vergleichs* [PISA 2003: Education Level of German adolescent: results of the second international comparison] Muenster: Waxmann.

Walpuski, M., Kampa, N., Kauertz, A., & Wellnitz, N. (2008). Evaluation der Bildungsstandards in den Naturwissenschaften. [Evaluation of educational standards in science]. *Der mathematische und naturwissenschaftliche Unterricht*, 61 (6), 323-326.

Walpuski, M. (2006). Optimierung von experimenteller Kleingruppenarbeit durch Strukturierungshilfen und Feedback. [Optimizing of work in small learning groups using structural guidelines and feedback]. Berlin, Logos Verlag. Studien zum Physik- und Chemielernen. Niedderer, Hans; Fischler, Helmut & Sumfleth, Elke.

ASSESSING STUDENTS' UNDERSTANDING OF LIGHT PROPAGATION AND VISIBILITY OF OBJECTS IN TWO DIFFERENT CONTEXTS

Hye Eun Chu
Nanyang Technological University, Singapore

David F Treagust
Curtin University of Technology, Australia

A.L. Chandrasegaran
Curtin University of Technology, Australia

Alexander Kauertz
University of Weingarten, Germany

Abstract

A large scale quantitative study involving 1,786 Years 7 -10 students was undertaken to evaluate their understanding of Light Propagation and Visibility of Objects, that is of two concept groups in two different contexts, using four pairs of two-tier multiple-choice items. Students' conceptual understanding in the Light Propagation concept group was found to be less context-dependent than in the Visibility of Objects concept group, with greater agreement between correct responses in the items relating to Light Propagation concepts. Also, these two concept groups were found to be moderately correlated, most likely as a result of students' limited experience in observing objects in complete darkness. Moreover, students' grade levels were found to have limited influence on their understanding of basic optics concepts. These research findings illustrate the need for teachers to provide students with a wide range of experiences in different contexts so as to improve their understanding of basic optics concepts.

Introduction

Several studies have highlighted students' alternative conceptions about optics concepts (Anderson & Karrqvist, 1983; Bengall, Goldberg, & Galili, 1983; Driver et al., 1994; Duit & Treagust, 1998; Galili & Hazon, 2000; Hey wood, 2005; Toh, Boo, & Woon, 1999; Shapiro, 1989). This study goes a step further and focuses, in particular, on elucidating the stability of and interrelationships between students' conceptions about light propagation and visibility of objects in different contexts across the first four years of secondary schooling. In recent years, the importance of the context in which science is learned has received more attention by researchers and curriculum developers (Nentwig & Waddington, 2005). Consequently, the purpose of the study was to demonstrate how students' understanding of optics concepts in different contexts may be conveniently evaluated using two-tier multiple-choice items and to provide evidence for the nature of context-based learning and its assessment.

Rationale

Using two-tier multiple choice items as formative assessment tools, teachers can identify the conceptions held by students that are not in agreement with scientific views. Based on these findings, relevant strategies may be formulated to challenge students' understandings in order to help them develop more scientifically acceptable views of science concepts (Treagust, 1995). The theoretical framework is based on the notion of students constructing their own knowledge in relation to their personal and classroom experiences (Scott et al., 2007).

Methods

Participants

The investigation was a large scale quantitative study involving 1,786 Korean students from Years 7 to 10. In Korea, students learn basic optics concepts in primary school (Years 3 – 6) and Year 8. The science taught is an integrated science curriculum from Years 3 to 10. After Year 10, students study science as four different subjects, physics, chemistry, biology and earth science. Students are required to choose subjects from these four. Years 7 to 9 are middle school and Years 10 to 12 are high school.

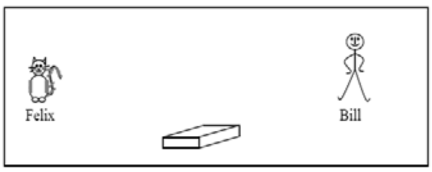
Research instrument

Data were collected based on the Light Propagation Diagnostic Instrument (LPDI). Students were given 30 minutes to complete the questionnaires. The LPDI was developed by the authors based on previous studies (Fetherstonhaugh & Treagust, 1992; Langley, Ronen & Eylon, 1997; La Rosa, Mayer, Patrizi & Vicentini-Missoni, 1984), and consisted of eight two-tier multiple-choice items. Each of four pairs of items investigated students' understanding of a particular concept in different contexts. Figure 1 shows an example of a paired item in the LPDI.

The first tier of each item evaluates students' content knowledge about a concept or phenomenon. The second tier requires students to justify the selection of their response in the first tier. The response to an item was considered to be correct if students answered both tiers of the item correctly. The eight items have been categorised in two concept groups, Light Propagation and the Visibility of Objects. The contexts of items in the two concept groups are summarised in Table 1.

The combined Cronbach alpha reliability of the eight items was 0.65, while the reliability of each concept group was 0.65 for Light Propagation concepts and 0.51 for Visibility of Objects concepts. According to Nunally and Bernstein (1994) a Cronbach alpha reliability coefficient greater than 0.7 indicates a high reliability, while values in the range 0.5 to 0.7 indicate moderate reliability. The reason for the low reliability of the Visibility of Objects concept group will be discussed later in the conclusion and implications section.

Item 7

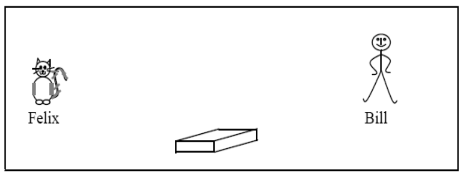


Felix the cat and Bill are in a completely dark room. There is no light in the room.
Felix the cat would:

A. not be able to see at all.
B. just be able to see the box.
C. see the box quite clearly.

The reason I chose my answer is because:
1. Light has to be reflected from the book to the cat's eyes.
2. Cats can see in the dark.
3. The cat is able to see objects by looking at them.
4. The cat will be able to see in the dark after adjusting its eyes to the darkness.

Item 8



This item is just like item 7. The room is still completely dark.
Bill would:
A. not be able to see the box at all.
B. just be able to see the box.
C. see the box quite clearly.

The reason I chose my answer is because:
1. We need light to be reflected to our eyes to be able to see in the dark.
2. People can just see in the dark.
3. We see by looking at objects.
4. We are able to see in the dark after our eyes have adjusted to the darkness.

Figure1. An example of one pair of items in Light Propagation Diagnostic Instrument (LPDI)

Data analysis

Item difficulties and discrimination indices were computed. The Kappa measure of agreement was considered to investigate the consistency in students' response in two different contexts in each paired items. The percentage of students' answers of the LPDI were analysed in two concept categories and the correlation between categories were considered. The SPSS (version 16) software was used for data analysis in this research.

Table 4. The contexts of the items in the Light Propagation Diagnostic Instrument

The contexts of items	
Light Propagation	Visibility of Objects
Item pair 1: Item 1 - During the day Item 2 - At night	**Item pair 2:** Item 3 - Non-luminous object Item 4 - Luminous object
Item pair 3: Item 5 - Observing lighted lamp from window above an obstruction Item 6 - Observing light propagation to windows above an obstruction	**Item pair 4:** Item 7 - Vision of cats in complete darkness Item 8 - Human vision in complete darkness

Results

Item Analysis

Agreement of students' answers in different contexts: The paired items (1 & 2; 3 & 4; 5 & 6; 7 & 8) each involved the same science concept in different contexts. The Kappa measure of agreement was used to show the consistency of students' responses in the two different contexts in each item pair. Students' correct answers were coded as 1 and wrong answers as 0. Therefore, the Kappa values indicated the students' consistency about correct and wrong answers in the paired items. A Kappa value of 0.5 represents moderate agreement, while a value above 0.7 represents good agreement (Peat, 2001). The two pairs of items (1 & 2, and 5 & 6) in the Light Propagation concept group involved the concept that 'light travels in straight lines in all directions until it strikes an object'. Each item pair showed Kappa values of 0.7 (Items 1 & 2) and 0.5 (Items 5 & 6). The two pair of items (3 & 4, and 7 & 8) in the Visibility of Objects concept group involved the concept that 'an object is visible because light is reflected from the object to the eyes'. Each item pair showed Kappa values of 0.4 (Items 3 & 4) and 0.3 (Items 7 & 8). Therefore, students' responses showed moderate to good agreement between the items in the item pairs relating to the Light Propagation concept group. However, there was lower agreement between the items in the item pairs relating to the Visibility of Objects concept group, suggesting that students' understanding about the concept of visibility presented in these items is more highly dependent on the contexts than students' understanding of light propagation.

The correlation between the two concept groups: The Pearson correlation coefficient value of r = 0.3 indicates that there was a significant correlation of medium strength between the two concept groups (Cohen, 1988). One reason for the limited correlation between the two concept groups is the different characteristics of the two concepts: while light propagation can be readily experienced by students, they have limited experience of being in complete darkness.

Students' general alternative conceptions

Light propagation: Four of the five general alternative conceptions identified were held by 13% - 30% of students in all school years in table 2. Furthermore, the alternative conceptions in the light propagation concept were not context-dependent, i.e. students displayed the same alternative conceptions in the two different contexts in each pair of items.

Students displayed the same alternative conception in Items 1 & 2 in the two different contexts of day and night when they suggested that "the light from a bulb comes out until it hits something, because light rays travel in a preferential way towards an object (D5)". Also, students showed the same alternative conception that light from lamp is visible at all points above an obstruction in Items 4 & 5 in two different contexts of observing a lighted lamp from windows above an obstruction and observing illuminated windows above an obstruction (A1). It means students' alternative conceptions in the two different contexts are not dependent on the context. Moreover, in item 1, 10% of students in Year 8 and 9 indicated the alternative conception that "light from the bulb stays on the light bulb, because light does not travel at all in the day time (A2)".

Table 2. Number of students with correct answer and the consistency of correct answer in each paired item

(% in the parentheses)

Year / Item	7 (n=410)	8 (n=458)	9 (n=367)	10 (n=551)	Total (N=1786)	Kappa value
1	188 (46)	216 (47)	138 (38)	278 (51)	820 (46)	0.7
2	191 (47)	227 (50)	149 (41)	302 (55)	869 (49)	
*1 & 2	161 (39)	182 (40)	123 (34)	258 (49)	724 (41)	
3	64 (16)	129 (28)	69 (19)	139 (25)	401 (23)	0.4
4	88 (22)	125 (27)	91 (25)	157 (29)	461 (26)	
*3 &4	34 (8)	67 (15)	42 (11)	89 (16)	232 (13)	
5	102 (25)	113 (25)	86 (23)	113 (21)	414 (23)	0.7
6	93 (23)	113 (25)	73 (20)	98 (18)	377 (21)	
*5&6	59 (14)	67 (15)	45 (12)	56 (10)	227 (13)	
7	72 (18)	95 (21)	114 (31)	109 (20)	390 (22)	0.3
8	226 (55)	262 (57)	220 (60)	278 (51)	986 (55)	
*7&8	61 (15)	86 (19)	110 (30)	100 (18)	357 (20)	

* The data refer to frequency of correct responses two both items in each pair

The visibility of objects: Five of the six general alternative conceptions identified were held by 11% - 35% of students in all school years are shown in Table 4. However, all the general alternative conceptions appeared in only one specific context in each pair of items.

Items 3 and 4 are in two different contexts, the visibility of a luminous object and a non-luminous object. In these items, students displayed a different alternative conception in each context by suggesting that "An opaque object (e.g. a flower) is visible because light is present around the object (D3)" and *"A person sees a lighted object (e.g. a candle flame) as bundles of rays radiating from the object (A1)".* In item 4, 11-15% of students from all school Years displayed an alternative conception about the light propagation diagram in the context of the luminous object (D4). These students chose the light propagating diagram that implies that light stays around the luminous object. However, this alternative conception of the diagram for light propagation was not displayed in the context of a non-luminous object. Also, items 3 and 4 are in two different contexts, vision of cats in complete darkness and human vision in complete darkness. In these items, students displayed a different alternative conception in each context by suggesting that "cats have the ability to see objects in the dark" and "people are able to see an object in a completely dark room after their eyes have adjusted to the darkness". It means students' alternative conceptions in the two different contexts are dependent on the context.

Moreover, of 12% students in Year 8 displayed an alternative conception in the context of an illuminated object when they suggested that "a person is able to see an opaque object because the object is located within the person's region of vision."

Table 3. Students' general alternative conceptions in the Light Propagation concept group

Alternative Conceptions	Choice Combination	Range of Percentage	School Grades			
			7	8	9	10
Light travels in a preferential way towards an object						
1. In the day time, the light from a bulb comes out until it hits something, because light rays travel in a preferential way towards an object.	Item 1(D5)	13-17	√	√	√	√
2. At night, the light from a bulb comes out until it hits something, because light rays travel in a preferential way towards an object.	Item 2(D5)	16-24	√	√	√	√
Light from a lamp is visible at all points above an obstruction						
4. We can see a lamp from a window above an obstructing wall because light from the lamp is visible at all points above the obstruction.	Item 5(A1)	24-30	√	√	√	√
5. All windows above an obstructing wall are illuminated because light from the lamp is visible at all points above the obstruction.	Item 6(A1)	26-30	√	√	√	√

Table 4. Students' general alternative conceptions in the Visibility of Objects concept group

Alternative Conceptions	Choice Combination	Range of Percentage	School Grades			
			7	8	9	10
Light stays around an object						
6. An opaque object (e.g. a flower) is visible because light is present around the object.	Item 3(D3)	15-21	√	√	√	√
7. A lighted object (e.g. candle flame) is emanating from the object and being received by eyes because light is present around object	Item 4(D4)	11-15	√	√	√	√
Objects can be seen without considering our eyes having to focus the rays on the retina						
9. A person sees a lighted object (e.g. a candle flame) as bundles of rays radiating from the object.	Item 4(A1)	16-24	√	√	√	√
10. Cats have the ability to see objects in the dark.	Item 7(C3)	26-35	√	√	√	√
11. People are able to see an object in a completely dark room after their eyes have adjusted to the darkness.	Item 8(B5)	14-22	√	√	√	√

Conclusions and Implications

Students' alternative conceptions in fundamental optics were investigated using two-tier multiple-choice items, LPDI (Light Propagation Diagnostic Instrument). The items consisted of two conceptual groups, Light Propagation,

and Visibility of Objects and each pair of items includes two different contexts. Students had to apply the same scientific concept in each pair of items in different contexts. Several conclusions may be drawn from this study.

First, students were more consistent in providing correct responses to the items in Light Propagation conceptual group compared in the Visibility of Objects conceptual group. Second, students displayed similar alternative conceptions across the school years as those that are reported in the research literature. Third, students generally could not apply their conceptions in basic optics in different contexts giving rise to several context-dependent alternative conceptions. Most students' alternative conceptions in the Light Propagation conceptual group were not context dependent while most students' alternative conceptions were context dependent in the Visibility of Objects concept group. Fourth, moderate correlations between the two concept groups were found in this research. Reasons for the moderate correlations may have been the low Kappa agreement and students' context dependent conceptions in the Visibility of Objects concept group. Also, the moderate correlations may have contributed to the low Cronbach's alpha reliability of the Visibility of Objects concept group.

This study has several pedagogical implications. First, the two-tier diagnostic test in basic optics has proven to be valuable diagnostic assessment tools that can help teachers diagnose their students' preconceptions before commencing their lessons. Second, the topic of optics should provide students with opportunities to compare concepts in different contexts in order to facilitate their conceptual development of the essential underlying concepts as well as being a useful means of formative assessment in classroom instruction and to institute remediation, where necessary. Third, it is essential that basic optics concepts are not taught in isolation but are progressively developed in higher grade levels. Lastly, further research is required to investigate how students improve their conceptual understanding in the context dependent concepts and context free concepts relating to the propagation of light and the visibility of objects.

References

Andersson, B., and Karrqvist, C. (1983). How Swedish pupils ages 12-15 years understand light and its properties. *European Journal of Science Education 5*, 387-402.

Bengall, S., Goldberg, F., and Galili, I. 1993. Prospective elementary teachers' prior knowledge about light. *Journal of Research in Science Teaching 30*, 1169-1187.

Cohen, J. W. (1988). Statistical power analysis for the behavioural sciences (2nd ed). Hillsdale, NJ: Lawrence Erlbaum Associates.

Driver, R., Squires, A., Rushworth, P., & Wood-Robinson, V. (1994). *Making sense of secondary science: Research into children's ideas.* London, UK: Rutledge.

Duit, R., & Treagust, D. F. (1998). Learning in science - from behaviorism towards social constructivism and beyond. In B. Fraser and K Tobin (Eds.), *International handbook of science education, Part 1.* (pp. 3-25). Dordrecht, The Netherlands: Kluwer.

Fetherstonhaugh, T., & Treagust, D. F. (1992). Students' understanding of light and its properties: Teaching to engender conceptual change. *Science Education*, *76*(6), 653-672.

Galili, I., and Hazon, A. 2000. Learners' knowledge in optics: Interpretation, structure, and analysis. *International Journal of Science Teaching, 22* (1), 57-88.

Heywood, D. S. 2005. Primary trainee teachers' learning and teaching about light : Some pedagogic implications for initial teacher training. *International Journal of Science Teaching, 27*(12), 1447-1475.

La Rosa, C., Mayer, M., Patrizi, P., & Vicentini-Missoni, M. (1984). Commonsense knowledge in optics: Preliminary results of an investigation into the properties of light. *European Journal of Science Education*, *6*(4), 387-397.

Langley, D., Ronen, M., & Eylon, B-S. (1997). Light propagation and visual patterns: Preinstruction learners' conceptions. *Journal of Research in Science Teaching, 34*(4), 399-424.

Nentwig, P. & Waddington, D. (Eds.). (2005). *Making it relevant: Context based learning of science.* Münster, Germany: Waxman.

Nunally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.), New York, NY: McGraw Hill.

Peat, J. (2001). *Health science research: A handbook of quantitative methods*. Sydney: Allen & Unwin.

Scott, P., Asoko, H. & Leach, J. (2007). Student conceptions and conceptual learning in science. In S. Abell and N. Ledermann (Eds.). Handbook of research on science education (pp. 31-56). Mahwah. NJ: Erlbaum.

Shapiro, B. L. (1989). What children bring to light: Giving high status to learners' views and actions in science. *Science Education, 73*(6), 711-733.

Toh, K. A., Boo, H. K., and Woon, T. L. 1999. Students' perspectives in understanding light and vision. *Educational Research* 41, no. 2: 155-162.

Treagust, D. F. (1995). Diagnostic assessment of students' science knowledge. In S. M. Glynn & R. Duit (Eds.), *Learning science in the schools: Research reforming practice* (Vol. 1, pp. 327-346). Mahwah, NJ: Lawrence Erlbaum.

IDEAS OF PRESCHOOL STUDENTS ON SOME ASTRONOMY CONCEPTS

Hüseyin Küçüközer
Balıkesir Üniversitesi

Ayberk Bostan
Balıkesir Üniversitesi

Abstract

The aim of this study is to determine preschool students' conceptions about the centre of the universe, the position of the stars during daytime and the brightest star at night. Although numerous studies have been conducted with preschool students about science education, few of them are on astronomy. 52 students (aged 6) from four different kindergartens in Balıkesir, Turkey were selected as the sample of the study. Research data were collected by a semi-structured individual interview protocol. The results of the study revealed that preschool students have some sort of ideas on each concept and most of these ideas are misconceptions. Individual interviews with preschool students demonstrated that students acquire these conceptions from their families, daily lives and observations.

Introduction

The theoretical framework of this study is based on the constructivist learning theory. Constructivism argues that students actively construct understanding from their experiences by using their already existing conceptual framework. Knowledge is constructed within certain social and material contexts in interaction with existing knowledge and new experiences or ideas. Furthermore, our knowledge about the world is not a mere copy of the reality outside us; it is our tentative construction about it. Scientific truth is not absolute but relative and so may change over time (Driver & Bell, 1986, Widodo, Duit & Müller, 2002). For researchers who subscribe to constructivism, students' existing knowledge and understanding are important because of their influence on the learning process. If students' prior knowledge is incorrect, they will construct wrong mental frames. These wrong ideas are generally called misconceptions or alternative frameworks and can be observed in all age groups and at all instruction levels. Thus, students may have various ideas about a subject before instruction.

So there are many studies on astronomy subjects but few of them are related to preschool students (Vosniadou and Brewer, 1992; Vosniadou and Brewer, 1994; Valanides, Gritsi, Kampeza, and Ravanis 2000; Hannust and Kikas, 2007). Vosniadou and Brewer (1994) worked with 20 students at the age of 6 years and 9 months. These students were at an elementary school in USA but were at the same age with our sample and did not receive instruction about the subject. The study examined children's concepts on the 'disappearance of the Sun at night', 'explanations of the day/night cycle', 'movement of the Moon' and 'disappearance of the stars during the day'. Preschool students had various ideas about the disappearance of the stars during the day. The most common misconceptions were 'the stars occluded by clouds', 'the stars move out into space', the stars go down on/in the ground', 'the stars move down, under the earth', 'the stars move down, unspecified as to which side of earth', 'the stars move somewhere else', 'the stars disappear' and 'stars stay where they are'. Furthermore, the study failed to categorize four students' responses. The findings demonstrate that only three students provided correct responses to this question. Researchers obtained the same ratio with regard to the responses to the other questions. In another study, Vosniadou and Brewer (1992) studied the same sample in terms of their mental models of the Earth's shape. Most first-grade students held a dual earth mental model or a mixed model and only three students gave correct responses. Other models about the Earth's shape included hollow sphere, flattened sphere and rectangular Earth. Valanides et al. (2000) studied with 33 preschool students in Greece what preschool students know about the shapes of the Sun and the Earth and how they perceive the concept of day-night cycle. Although most of the students stated before and after the instruction that the Sun and the Earth's shapes are sphere-like, few described their shapes as hemisphere, disk, cube or pyramid. Hannust and Kikas (2007) worked with 113 preschool students in Estonia and investigated the concepts of gravity and the shape of the Earth before and after instruction. They assigned preschool students to control and experiment groups. Before the instruction, 48% of the students in the experiment group and 58% of students in the control group provided correct responses to the question about the Earth's shape. After the instruction, 88% of the students in the experiment group gave correct answers, while the ratio of correct answers was 81% for the control group. These results suggest that nearly half of the preschool students had misconceptions about shape of the Earth before instruction, whereas most of the students provided correct responses to the question following the instruction. These studies demonstrate that preschool students have some ideas about astronomy concepts but most of them are misconceptions.

It should be noted that the student sample had received no formal instruction before, so the students' ideas were their naive ideas. Naive ideas are defined by perception of surrounding (Marin, 2000). Naive ideas are not affected by instruction.

Rationale

This study aimed to investigate the ideas of preschool students about the centre of the universe, the position of the stars during the day and the brightest star at night.

A literature review (Pfundt and Duit, 2007) reveals that there are few studies on preschool students' ideas about the disappearance of the Sun at night, explanations of the day/night cycle, movement of the Moon, disappearance of the stars during the day, shape of the Sun and the Earth and gravity. To the best of our knowledge, no study has been carried out with preschool students on the concepts of 'the centre of the universe' and 'the brightest star at night'. Therefore, this study will, for the first time, reveal the ideas of preschool students about these conceptions. Yet, Vosniadou and Brewer (1994) investigated preschool students' ideas about the disappearance of the stars during the day. This is the first study conducted in Turkey to examine preschool students' ideas about the disappearance of the stars during the day.

Methods

52 students in the sample group of the study were selected from four different kindergartens located in the central province of Balıkesir (Turkey) and all of the students were six years old. None of the students received any instruction about the research questions.

To explore preschool students' ideas about astronomy concepts, semi-structured individual interviews were carried out for about 20 minutes. While preparing the interview questions about the events on which students can do self-observation, some of the studies in the literature were benefited (Baxter, 1991; Valanides, et al., 2000; Hannust and Kikas 2007). After preparation, the questions were checked by one science education expert and five physics education experts for content validity. Before the research process, necessary corrections were made in accordance with their recommendations. As for the study content, the following questions were addressed to the preschool students: 'Is there a centre of the universe? If there is, where could it be?', 'When we look the sky we cannot see the stars during the day. Why do you think it is?' and 'When you look the sky at night you see many stars. What is the brightest one called?'. The questions were reformulated when the preschool students failed to understand the questions. The students' responses to each question were analyzed to determine their conceptions about astronomical phenomena. Similar explanations were classified under the same categories. For example, seventeen students were assigned to the Settlement location category. The student responses included residential areas such as Balıkesir, Izmir, Istanbul, Ankara (cities in Turkey), the Black Sea, the city centre and houses we live in, supermarkets where we buy our food and schools etc.

The frequency values of these classified explanations are presented in tables for each question in the results section. Under the correct answer in the first row, the tables show the answers that can be regarded as misconceptions and the answers that cannot be coded by researchers.

Results

Centre of the Universe

In Table 1, the ideas of preschool students on the centre of the universe as revealed by the interviews are summarized in the subsequent sections.

Table 1. Students' ideas about the centre of the universe

Ideas	**Number of Students f (%)**
No centre	13 (25)
Settlement location	17 (32.69)
The Sun	7 (13.46)
The Earth	4 (7.69)
Planets	3 (5.77)
Space station	3 (5.77)
Uncodable	3 (5.77)
No explanation	2 (3.85)

%25 of all students gave correct responses to this question by pointing out that there is no centre of the universe. This rate is quite high. During the interviews, these students stated that they had never heard about the centre of the universe, so there is no centre of the universe. An example of the preschool students' responses is as follows: "If there were a center of the universe, we would have heard it". 13.46% of all the students believed that the centre of the universe as it is very shiny. Küçüközer (2007), Lemmer, Lemmer and Smith (2003), and Trumper (2000) and Trumper (2001, b) found in their studies that university students held the misconception that the Sun is the centre of the universe. 69% of all the students believed that the Earth is the centre of the universe. Student 17's explanation is given below as an example for this category.

Interviewer: Where do you think the centre of the universe is?

Student 17: Earth is centre of the universe.

Interviewer: Why?

Student 17: Because we live on the Earth and all living things live on the Earth.

32.69% of all the students mentioned settlement location as the centre of the universe. In this answer category, the following locations were said to be the centre of the universe: Istanbul by two students, Ankara by one student, Izmir by one student, the Black Sea by one student, the city centre by two students, and the rainbow by one student. Below is an example of the response category of settlement location.

Interviewer: Where do you think the centre of the universe is?

Student 9: I think Istanbul is the centre of the universe.

Interviewer: Why do you think Istanbul is the centre of the universe?

Student 9: Because my aunt lives there and we often go there. Istanbul is a very big city so it is the centre of the universe.

Three students' responses could not be categorized and two students failed to provide any response to this question.

Stars

In Table 2, preschool students' ideas about the position of the stars during the day as revealed by the interviews are summarized in the subsequent sections.

Table 2. Students' ideas about the position of the stars during the day

Ideas	**Number of Students f (%)**
Sunlight	19 (36.54)
Clouds	14 (26.92)
Movement of the stars	14 (26.92)
God	2 (3.85)
Colour of the stars	2 (3.85)
No explanation	1 (1.92)

36.54% of all the students correctly answered the question concerning the position the stars during the day. Their explanation was that during daytime, the sunlight is so bright that it prevents us from seeing the stars and their position. The most common misconception is the category of clouds. 26.92% of all the students stated that passing clouds cover the stars during daytime or the stars are too high that they stay behind the clouds. Vosniadou and Brewer (1994) reported similar misconceptions and three students in their study stated that the stars are behind the clouds during daytime. An example of such a response is given below.

Interviewer: We cannot see the stars during day time. Why can't we see them?

Student 5: Because of clouds.

Interviewer: Can you explain your answer? How do clouds prevent seeing the stars?

Student 5: During daytime clouds cover the sky and stars remain behind the clouds.

Another misconception relates to the movement of the stars. The responses in this category are explained as follows: seven students said that the stars move elsewhere during the night, five students said that the stars move into the space during daytime and two students said that the stars move with the Moon to another place at night. In Vosniadou and Brewer's (1994) study, there were five different misconceptions in the movement of the stars category. For example, two students said that 'the stars move out into space', one student said that 'the stars go down on/in the ground', three students said that 'the stars move down, under the earth', two students said that 'the stars move down, unspecified as to which side of earth', and one student said that 'the stars move somewhere else'. All these misconceptions are similar to our findings. Below is an example in 'the stars move somewhere else' category.

Interviewer: We cannot see the stars during day time. Why can't we see them?

Student 23: Because stars move to other places.

Interviewer: Where do the stars move?

Student 23: They move to other places when night falls.

Two students thought that the stars cannot be seen during the day since God prevents it and mentioned their families as the origin of their ideas. Brickhouse, Dagher, Letts and Shipman (2000) stated that students are influenced by their religious beliefs when they explain scientific concepts and Küçüközer (2007) also found similar results among preservice science teachers. Thus, students' religious beliefs are very effective on their ideas in all age groups. Two students said that during daytime, the sky and the stars are the same colour so stars cannot be seen. Only one student did not provide any response to this question.

Brightest Star at Night

In Table 3, the preschool students' ideas about the brightest star at night as revealed by the interviews are summarized in the subsequent sections.

Table 3. Students' ideas about the brightest star at night

Ideas	**Number of Students f (%)**
Shooting star	14 (26.92)
Moon	13 (25)
North Star (Polaris)	4 (7.69)
Shepherd's star	3 (5.77)
Uncodable	9 (17.31)
No explanation	9 (17.31)

Preschool students gave several answers about the brightest star at night but there was no correct answer among them. This was not a surprising result because students did not receive any instruction about the subject. The response of the shooting star was the most common misconception among the students. They saw the stars as falling and made a wish. 25% of all the students thought that the Moon, the biggest celestial body, as the brightest star. The symbol of the moon and star on our country flag influenced the students' ideas and they mentioned their families as the source of their ideas. Below is an example for this category.

Interviewer: Which star is the brightest star at night?

Student 41: Moon is the brightest star.

Interviewer: How did you learn this?

Student 41: I saw it on our flag.

In their study with elementary school students, Küçüközer et al. (2009) reported misconceptions about the Moon, the North star and Shepherd's star. This result shows that similar misconceptions could be observed in different age groups. From the above information, it is clear that everyday observations affect preschool students' ideas. Four students mentioned the North Star and three students mentioned Shepherd's star as the brightest star. 17.31% of all students' responses could not be categorized and again 17.31% of all the students failed to provide any response to this question. The rate of the uncodable responses and no explanation categories was higher than that of other responses.

Conclusions and Implications

The preschool students' responses demonstrate that they have various ideas about the centre of the universe, the position of the stars during the day and brightest star at night. Since preschool students did not receive any instruction on the subject, their ideas are naive ideas and are not affected by instruction. Not many of these naive ideas are compatible with scientific facts. The results reveal that most of the preschool students' ideas are misconceptions. According to the research data, the preschool students' misconceptions are very diverse. The most common misconceptions according to the survey results are as follows. As for question about the position of the stars during daytime, the most common responses with misconceptions are clouds and movement of the stars. In response to the question about the brightest star at night, the 'Moon' was the most common misconception. To the question about the centre of the universe, the students' responses mentioned settlement location and the Sun misconceptions. Different studies with high school and university students also came up with a similar misconception that Sun is the centre of the universe (Trumper, 2000; Trumper, 2001,a; Trumper, 2001,b; Küçüközer, 2007; Lemmer et al., 2003). The misconception that the Sun is the centre of the universe was observed among preschool students, as well as among high school and university students. Studies with different age groups also reported similar misconceptions.

Misconceptions start to be observed in the early years and continue until later years (Gülçiçek and Yağbasan, 2003). Preschool students pointed out during the interviews that their ideas are affected by self-observations, their experiences and their families' contributions. Since preschool students did not receive any instruction on astronomical subjects, the conceptions they developed are based on their perceptions about their environment. The responses of the preschool students to the interview questions reflect how they perceive the events around them. In their explanations, they clarified how they perceive events and they rationalized their accounts by their observations. For example, the students' responses to the question about the brightest star at night were influenced by their observations. And in the question about the centre of the universe, the students' responses were affected by their everyday experiences. The results of this study could be taken into consideration when an instructional context is constructed for preschool students. Studies conducted to examine preschool students' ideas are very few. Researches studying in this field may carry out studies to determine preschool students' ideas on different topics. They can also investigate the source of preschool students' ideas and organize instructional activities to eliminate such misconceptions.

References

Baxter, J. (1991). A Constructivist Approach to Astronomy in the National Curriculum. *Physics Education,* 26, 38-45.

Brickhouse, N.W., Dagher, Z.R, Letts, W.J. & Shipman, H.L. (2000). Diversity of Students' Views about Evidence, Theory and the Interface between Science and Religion in an Astronomy Course. *Journal of Research in Science Teaching,* 37(4), 340-362.

Dove, J. (2002). Does The Man In The Moon Ever Sleep? An Analysis of Students Answers about Simple astronomical Events: A Case Study. *International Journal of Science Education,* 24(8), 823-834.

Driver, R. & Bell, B. (1986). Students' Thinking and Learning of Science: A Constructivist View. *School Science Review*, 67(240), 443-456.

Gülçiçek, Ç. & Yağbasan, R. (2003). Fen Eğitiminde Kavram Yanılgılarının Karakteristiklerinin Tanımlanması. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*, 1(13), 102-120.

H. Pfundt & Duit, R., 2007. Bibliography STCSE Full Version. http://www.ipn.uni-kiel.de/aktuell/stcse/stcse.html.

Hannust, T. & Kikas, E. (2007). Children's Knowledge of Astronomy and Its Change In Course Of Learning. *Early Childhood Research Quarterly.* 22, 89-104.

Küçüközer, H. (2007). Prospective Science Teachers' Conceptions about Astronomical Subjects. *Science Education International,* 18(1), 113-130.

Küçüközer, H., Korkusuz, M.E., Küçüközer, A. & Yürümezoğlu, K. (2009). The Effect of 3D Computer modeling and Observation-Based Instruction on the Conceptual Change Regarding Basic Concepts of Astronomy in Elementary School Students, *Astronomy Education Review,* 8(1).

Lemmer, M., Lemmer, T. N. & Smit, J. J. A. (2003). South African Students Views of the Universe. *International Journal of Science Education*, 25(5),563-582.

Marin, N. (2000). What Is the Relationship Between Social Constructivism and Piagetian Constructivism? An Analysis of the Characteristics of the Ideas within Both Theories. *International Journal of Science Education*, 22(3), 225-238.

Trumper, R. (2000). University Students' Conceptions of Basic Astronomy Concepts. *Physics Education,* 35(1), 9-15.

Trumper, R. (2001,a). A Cross-age Study of Senior High School Students' Conceptions of Basic Astronomy Concepts. *Research in Science &Technological Education,* 19(1), 97-107.

Trumper, R. (2001,b). A Cross-College Age Study of Science and Nonscience Students' Conceptions of Basic Astronomy Concepts in Preservice Training For High-School Teachers. *Journal of Science Education and Technology,* 10(2), 189-195.

Valanides, N., Gritsi, F., Kampeza, M. & Ravanis, K. (2000). Changing Pre-school Children's Conceptions of the Day/Night Cycle. *International Journal of Early Years Education,* 8(1).

Vosniadou, S. & Brewer, W.F. (1992). Mental Models of the Earth: A Study of Conceptual Change in Childhood. *Cognitive Psychology*, 24, 535-585.

Vosniadou, S. & Brewer, W.F. (1994). Mental Models of the Day/Night Cycle. *Cognitive Science*, 18, 123-184.

Widodo, A.; Duit, R. & Müller, C. (2002). Constructivist Views of Teaching and Learning in Practice: Teachers' Views and Classroom Behaviour. Annual Meeting of the National Association for Research in Science Teaching, New Orleans.

A LEARNING PROGRESSION FOR ELEMENTARY SCHOOL STUDENTS' UNDERSTANDING OF PLANT NUTRITION[1,2]

Alicia C. Alonzo
Michigan State University (USA)

Matthew Benus, William Bennett, & Brian Pinney
University of Iowa (USA)

Abstract

Learning progressions – descriptions of increasingly sophisticated ways of thinking about or understanding a topic (National Research Council, 2007) – represent a promising framework for bridging the chasm between cognitive science research and methods for teaching and assessing science. Learning progressions rely upon cognitive science research to describe a path (or set of paths) which students might take in moving from novice to expert understanding. They have been promoted as solutions to a wide range of current educational issues; however, much work remains to be done before learning progressions could be used in envisioned ways. There are still very few examples of empirically-validated learning progressions in science, particularly at the elementary school level. This paper presents a tentative learning progression for describing elementary school students' developing understanding of plant nutrition, a key topic in the elementary school science curriculum. A review of both standards documents and literature on student thinking about plant nutrition was used to develop a preliminary version of the learning progression. An interview protocol was developed to assess student understanding relative to this learning progression, and video-taped interviews with 30 students in grades 1-6 were coded using the preliminary learning progression. This analysis was used to revise the learning progression, resulting in the version presented here.

Introduction and Rationale

Learning progressions – descriptions of increasingly sophisticated ways of thinking about or understanding a topic (National Research Council [NRC], 2007) – have generated much interest in the United States (US) science education community. The US National Science Foundation (NSF) has issued special calls for projects focused on learning progressions (NSF, 2005, 2006, 2009); the US National Research Council has advocated for their development and use to inform science curricula (NRC, 2007) and assessments (NRC, 2006); and the most recent framework for the science portion of the US National Assessment of Educational Progress (NAEP) called for the test to include learning-progression-based items (National Assessment Governing Board, 2006). Learning progressions have been promoted as solutions to a wide range of current educational issues, including a lack of curricular coherence (Gotwals & Songer, 2008), developmental inappropriateness of curricula (Stevens, Shin, Delgado, Krajcik, & Pellegrino, 2007), misalignment between instruction and assessment (Black & Wilson, 2007), and weaknesses in support for valued teaching practices (Alonzo, 2009; Furtak, 2009). Learning progressions are viewed as a means of bridging the chasm between cognitive science research and methods for teaching and assessing science. US science educators hope that learning progressions will inform research-based approaches to the development of standards and curricula. By laying out a pathway for the way in which learning develops over a

[1] The work reported in this paper was supported by a grant from the Iowa Mathematics & Science Education Partnership (Grant # IMSEP1-10-08) and a research assistantship awarded by the University of Iowa's Department of Teaching & Learning.
[2] An expanded version of this paper is available upon request from the first author at alonzo@msu.edu.

long period of time, standards and curricula might be designed so that students' science learning builds from year to year, resulting in more durable and deep understandings. In the classroom, learning progressions may serve as a framework to focus greater attention on student thinking. By laying out a series of levels students might move through in coming to fully understand a given concept, learning progressions may provide important supports for formative assessment practices, helping teachers to recognize their students' "wrong" ideas as building blocks on the way to full understanding and to use these ideas to inform decisions about instructional next steps.

Despite these promises, much work remains to be done before learning progressions could be used in envisioned ways. There are still very few examples of empirically-validated learning progressions in science, particularly at the elementary school level. This is problematic for at least two reasons:

1) For both classroom and large-scale applications, learning progressions describing student thinking across a significant proportion of the curriculum would be required in order for these tools to significantly impact science teaching and learning.
2) There are limited examples illustrating the process of defining, assessing, and validating learning progressions. Such examples are important to the research community, which is still grappling with a number of issues related to learning progressions and the nature of student thinking they try to capture.

This paper describes the development and refinement of a learning progression describing elementary school students' thinking about plant nutrition. This is a key topic in the elementary school curriculum – both in terms of the amount of time US students spend learning about plants and in terms of its conceptual significance. The ability of plants to harness energy from the Sun to create their own food is vital to the existence of life on Earth. Knowledge of this process is a precursor to developing full understanding of ecological systems and carbon transformation. Both of these more advanced topics are critical for informed environmental decision-making.

The paper starts with a brief discussion of the construct of learning progressions. This is followed by a description of the two-step process we used to develop and refine a preliminary learning progression for elementary school students' understanding of plant nutrition. The resulting learning progression is then presented.[1] The paper concludes with a discussion of the learning progression and implications for future work.

Learning Progressions

Learning progressions grew out of earlier work by Mark Wilson and colleagues (e.g., Roberts, Wilson, & Draney, 1997) with "construct maps" whose defining features are 1) "a coherent and substantive definition for the content of the construct" and 2) "an underlying continuum" (Wilson, 2005, p. 26). However, while construct maps are "agnostic" with respect to the theoretical underpinnings of the proposed continuum, learning progressions are firmly grounded in research on how students learn a particular concept.

Since the "learning progression" construct is relatively new, consensus as to an exact format and essential features remains somewhat elusive. Examples of many different types of learning progressions exist in the literature (e.g., Alonzo & Steedle, 2009; Mohan, Chen, & Anderson, 2009; Schwarz, Reiser, Davis, Kenyon, Acher, Fortus, et al., 2009; Shavelson, Yin, Furtak, Ruiz-Primo, Ayala, Young, et al., 2008; Smith, Wiser, Anderson, & Krajcik, 2006; Songer, Kelcey, & Gotwals, 2009). However, most examples share some important features:

1) containing ordered descriptions of student thinking – rather than of content;
2) describing student thinking in the "messy middle" – rather than just the beginning (naïve conceptions) and end (scientifically correct conceptions).

[1] Due to space limitations, the interview protocol used to elicit students' ideas about plant nutrition and illustrative examples from student interviews have not been included here but are available in the expanded version of the paper.

The Plant Nutrition learning progression described in this paper shares these features. It spans a relatively long period of time – elementary school (grades K-6) – although one might expect students to progress through many of its levels during a well-designed curriculum unit. As described below, it posits 6 levels of student thinking, across three different progress variables.

Methods

Developing a Preliminary Learning Progression

The development of a learning progression is necessarily an iterative endeavor, involving successive refinements of hypotheses about the way in which learning progresses. Briggs, Alonzo, Schwab, and Wilson (2006) lay out a process for initial development of a learning progression. In this approach, the first step is defining the top level of the learning progression. As the top level can be viewed as representing targeted understandings, it is defined through a review of standards documents. For the Plant Nutrition learning progression, we consulted national standards documents (NRC, 1996; American Association for the Advancement of Science, 1993), as well as curriculum materials for teacher education (LessonLab Research Institute, 2008).

At the top level, students are expected to understand the process (photosynthesis) by which plants make their own food, using light energy and water and carbon dioxide. They are expected to know that plants do not take in any other types of food, although they require additional nutrients for growth. The food they produce supplies energy and raw materials for growth; it may also be used by animals (who cannot make their own food). Plants are unique in their ability to transform light energy into chemical energy (in food). Examination of the content contained in the top level resulted in the identification of three related ideas: *Food for Plants*, *Energy for Plants*, and *Plants as Producers*. Thus, the top level description was split into descriptions for the top levels of three progress variables. The *Food for Plants* and *Energy for Plants* progress variables pertain to the source of plants' food and energy, respectively. *Plants as Producers* deals with the relationship between plants and animals, especially with respect to food.

To inform development of the lower levels of the learning progression, a review of literature on student thinking about plant nutrition (e.g., Anderson, Sheldon, & Dubay, 1990; Barker & Carr, 1989; Barman, Stein, McNair, & Barman, 2006; Bell, 1985; Cañal, 1999; Çepni, Taş, & Köse, 2006; Leach, Driver, Scott, & Wood-Robinson, 1996; Marmaroti & Galanpoulou, 2006; Özay & Öztas, 2003; Smith & Anderson, 1984) was undertaken. This literature base permitted us to identify common ways that young children think about each of the three progress variables. Data – including ages at which different ideas are prevalent – as well as a logic-based process was used to order ideas within each strand and to create levels which describe how a student might be expected to think about plant nutrition across the three progress variables.

Developing an Interview Protocol to Assess Students' Understanding Relative to the Preliminary Learning Progression

In order to evaluate the extent to which the preliminary learning progression matched the understandings expressed by students, we designed an interview protocol to capture student ideas about plant nutrition. Using the initial learning progression, questions were identified from the literature and additional questions were written to elicit students' understandings with respect to each progress variable. One challenging aspect of this development process was the need to ask questions which could be answered by students at all levels of the learning progression. This requirement led to a branching interview protocol, in which follow-up questions, based upon students' earlier responses, were critical to our ability to engage students in a conversation which revealed their understanding of plant nutrition. In addition, we started the conversation with questions which – although not directly related to plant nutrition – allowed all students to share their ideas about plants. Responses to these questions provided the basis for later interview questions, which more directly explored students' ideas relative to the Plant Nutrition learning progression.

Gathering Evidence to Support Revision of the Preliminary Learning Progression

The interview protocol was used to elicit evidence of student thinking relative to the preliminary Plant Nutrition learning progression. Thirty students in grades 1-6 (approximately 6-12 years of age) participated in this phase of the research. The students all attended an elementary school in a suburb in the Midwestern US. Each student was interviewed individually in a quiet space in the school – the cafeteria when not in use or an empty classroom or office. All interviews were conducted by the first author. The interviews lasted between 5 and 23 minutes, with an average length of 11.8 minutes (SD = 4.0).

Interpreting Evidence to Support Revision of the Preliminary Learning Progression

The preliminary learning progression was used to code the interview transcripts. In this process, we sought evidence for students' understanding relative to the three progress variables. There was one code for each level of the progress variable, plus one code to flag student thinking which was not included in the existing learning progression but which seemed related to students' understanding of plant nutrition. Codes were applied only to segments of the interview which were identified as related to students' understanding of plant nutrition. Each segment was defined to include a single student idea (including the interviewer questions as necessary to make sense of the student talk) and could be assigned multiple codes. Four researchers (the authors) independently coded all of the interview transcripts and met to reach consensus about a final set of codes.

Revisions to the learning progression were based upon analysis of the codes, with particular attention to student ideas which were not adequately captured by the learning progression.

Results

In this section, we present the revised learning progression, in its current form. It is important to note that this is still a relatively un-tested product. Although research literature was used to develop the learning progression, and student interview data was used to refine it, the iterative process of refining a learning progression necessarily requires multiple cycles – with revisions to the learning progression and associated assessments in each cycle. (See Alonzo & Steedle (2009) for additional discussion of the revision process.) The learning progression presented here is at the beginning of these cycles of revision. Next steps for testing and revising the learning progression are discussed below.

Food for Plants

This progress variable (shown in Table 1) is concerned with the source of food for plants. As such, it depends critically upon students' own definitions for "food". At the lowest level (1), students are likely to have non-scientific ideas about this word. They may view everything that plants (and other organisms) need to live and grow as "food". Alternatively, at this level, students may not think that plants need "food" at all, perhaps viewing food as things like pizza and hamburgers that are eaten by people. At the next level (2), students hold the very typical idea that plants get their "food" from the environment. In interviews, students at this level often expressed the idea that plants take in food through their roots and identified a variety of different substances as being food for the plant. The next level (3) represents a transition between the Level 2 idea that plants get "food" from their environment and ideas at higher levels of the learning progression, at which students recognize that plants make their own food. Students at Level 3 think that food for plants consists of substances from their environment mixed together or transformed. At this level, students do not assign an active role to plants in the mixing or transformation process.

Table 1. *Food for Plants* Progress Variable of the Plant Nutrition Learning Progression.

Level	Description
1	• Everything that plants take in to live and grow is considered to be "food" OR • Plants are not considered to need "food".
2	• Plants get things they need to live and grow (including "food") from their environment.
3	• Plants' food consists of things from their environment (such as water, minerals from the soil) mixed together or transformed. *NOTE: This is different from level 4, in which plants have a more active role in the mixing/transformation process.*
4	• Plants produce their own food, using things from their environment. *NOTE: At this level, students may not know what plants use to make food or they may have incorrect ideas about the things which plants use to make food.*
5	• Plants produce all of their own food using water, carbon dioxide, and light. • Plants also take in other things from the environment to live and grow, but these things are not food.
6	• Plants produce their own food (in the form of glucose or sugar) using water, carbon dioxide, and light. Oxygen is produced as part of this process, which is called photosynthesis. • Photosynthesis isn't enough for plants to live and grow; plants also need other (non-food) things (such as minerals)

Level 4 is the first level at which students recognize that plants actually produce their own food; however, they may hold incorrect ideas about the substances that plants use in this process. Students may not have abandoned their earlier ideas about plants taking in food from their environment, thinking that plants supplement the food that they produce with additional food from their environment. Level 5 represents another important transition: the recognition that plants produce <u>all</u> of their own food. In addition, at this level, students are able to accurate identify the "inputs" that plants need in order to produce their own food (water, carbon dioxide, and light). Finally, Level 6 represents full understanding of the *Food for Plants* progress variable. In addition to the understandings at Level 5, students at this level can identify the type of food that plants produce (glucose, or sugar) and recognize that oxygen is produced as part of this process. If students mention photosynthesis, they are able to connect this scientific term to the production of food. In addition, at this level, students can differentiate between substances that plants need for photosynthesis and those needed for other life-sustaining processes.

Energy for Plants

This progress variable (shown in Table 2) is concerned with the source of energy for plants. Similar to the previous progress variable, this one depends critically upon students' own definitions for "energy". Strategically, this progress variable includes students' understanding of the role of the sun/light in plant growth, as this was thought to be a way for students with limited understanding of the word "energy" to still engage with ideas relevant to this progress variable. At the lowest level (1) students may not think that plants need the sun/light to live and grow or think that the sun is important for plants' survival for reasons other than energy, food, or growth. At this level, students associate energy with movement. Thus, they do not think that plants have energy <u>or</u> they think that plants use energy to move. At the next highest level (2), students recognize that plants need light/sun to grow and that plants have energy; however, their ideas about both light/sun and energy are vague. They cannot explain how the sun/light helps plants to grow or identify the source of energy for plants. Students at this level may retain some ideas about energy being related to movement, thinking that plants use energy to move parts of the plant or substances (such as food). At Level 3, students still have not developed an explanation for why plants need sun/light to grow. While they now recognize that plants get energy from their environment, they have not yet connected sun/light and energy. Students at this level think that plants use energy to grow, although they can provide only simple, unelaborated statements about this idea.

Table 2. *Energy for Plants* Progress Variable of the Plant Nutrition Learning Progression.

Level	Description
1	• Plants may not need sun/light to live. • Plants need sun/light for reasons other than energy/food/growth (e.g., warmth, source of vitamins). • Energy is associated with movement – either plants do not have energy or they use energy to move.
2	• Plants need sun/light to grow. • Plants have energy, but its source is unknown. • Plants use energy to move substances (e.g., food) or to stay healthy.
3	• Plants need sun/light to grow. • Plants get energy from things they take in from the environment (not including sun/light). • Plants use energy to grow.
4	• Plants need sun/light to make food or because it provides energy. • Plants get energy from things they take in from the environment (including sun/light and at least one other thing) <u>or</u> plants get energy from food that they produce. • Plants use energy to make food.
5	• Plants need sun/light because it is their only source of energy. Sun/light provides energy for the plant to make its own food. • Plants store energy, but the way this happens is unknown.
6	• Plants transform energy from sun/light into energy in food. • Plants get energy from sun/light and from the (stored) food that they produce.

At the next level (4), students are able to associate plants' need for sun/light with either the production of food or a need for energy. However, students at this level may view sunlight as just one of several components which plants mix together to produce food. They identify sun/light as a source of energy for plants, but also think that plants get energy from other environmental sources. At this level, students may recognize that plants use energy to make food; however, they may not explicitly connect this with sun/light. At Level 5, students recognize that sun/light is the only source of energy for plants and that plants use this energy to produce their own food. The idea of energy storage is introduced at this level, although students do not yet understand how this occurs. Finally, Level 6 represents full understanding of the *Energy for Plants* progress variable. In addition to the understandings at Level 5, students at this level understand that plants transform energy from sun/light into energy in food. They recognize that the energy plants use comes from the (stored) food that they produce.

Plants as Producers

This progress variable (shown in Table 3) is concerned with the relationship between plants and animals that exists because of plants' unique process of photosynthesis. While the progress variable focuses primarily on plants' production of food which can be eaten by humans and other animals, it also includes ideas about plant's independence from people and gas exchange between people and plants. Students' ideas in this progress variable may be particularly influenced by the context in which they are living and learning. For example, current environmental concerns in the US (and other countries) may lead to students' increased awareness about the gas exchange aspects of this progress variable, as compared to their counterparts just a few years ago.

At the lowest level (1), students view plants as incapable of living without human intervention. For example, students might think that people need to "feed" plants, by supplying them with fertilizer: At this level, students may not view plants as having any role in the growth of people/animals. For example, they may not recognize that people/animals can eat plants.

At the next level (2), students make the important leap to recognize that – under the right conditions – plants can grow independently from other living things. Students at this level recognize that plants can get the things they need to live and grow from their environment. At this level, students recognize that people/animals eat plants for food. However, they do not understand that people depend upon plants for food (because they eat not only plants, but animals which eat plants). During the interview, students were usually asked if people could still eat pizza and hamburgers if there weren't any plants; students at Level 2 answered this question in the affirmative or expressed the view that people could still eat these types of foods but that they would not be as healthy. Students at this level also have ideas related to plants' role in gas exchange – either that people need plants in order to breathe or that people need plants because they help the environment in some way. These ideas are still somewhat under-developed, but seem to be important precursors to the more detailed understandings at higher levels At Level 3, students move beyond a recognition that people eat plants to the idea that plants actually produce food for people. (At this level, students may still be referring to parts of the plant, such as fruits or vegetables, when stating that plants "produce" food. Or they may have a "mixing" or "transforming" idea of "production", similar to Level 3 of *Food for Plants*.) Students at this level also understand that people depend upon plants not only directly for food, but because we eat animals which eat plants. However, as at Level 2, there is not yet a recognition that this means people could not survive without plants; students may continue to think that people/animals simply wouldn't be as healthy without plants. Students' ideas about gas exchange also become a bit more refined at this level; they identify either that plants provide a good kind of air for people or that plants take away kind of air that is bad for people.

Table 3. *Plants as Producers* Progress Variable of the Plant Nutrition Learning Progression.

Level	Description
1	• Plants require human intervention to grow. • Plants may or may not help people/animals to live and grow.
2	• Plants can grow independent from other living things because – under the right conditions – it is possible for them to get things they need to live and grow from their environment. • People/animals eat plants for food. • People need plants in order to breathe and/or to help the environment (e.g., reducing global warming).
3	• Plants can grow independent from other living things because they get the things they need to live and grow from their environment. • Plants produce food for people. Plants do something with the food that they take in from the environment (before it can be used by people/animals). • People eat plants and animals which eat plants. • Plants either provide a good kind of air for people or take away a kind of air that is bad for people.
4	• Plants can grow independent from other living things because they produce their own food. Plants may use additional food from their environment. • People cannot survive without plants because people eat both plants and animals which eat plants. • Plants provide a good kind of air for people and take away a kind of air that is bad for people.
5	• Plants can grow independent from other living things because they produce all of their own food. • Plants are unique among living things because they can make their own food; the food that plants produce is the ultimate source for all living things on Earth. • Plants use a component of air that people breathe out and produce a component of air that people breathe in
6	• Plants can grow independent from other living things because they produce all of their own food, in the form of glucose (sugar). • Photosynthesis allows plants to use energy from the sun to create food; this is the food source for all living things on Earth. • During photosynthesis, plants use carbon dioxide that people breathe out and produce oxygen that people breathe in. Plants also require oxygen and release carbon dioxide.

Level 4 is dependent upon students' understanding that plants produce their own food. Students cite this as the reason that plants can grow independent from other living things, although they still may think that plants take in additional food from the environment. Building directly from the Level 3 idea that people eat animals which eat plants, students at Level 4 recognize that people cannot survive without plants. Students at this level have also expanded their Level 3 ideas about gas exchange to recognize that plants both provide a good kind of air for people and take away a kind of air that is bad for people. At Level 5, students recognize that plants can grow independent from other living things because they produce all of their own food. This level includes an important advance over the lower levels: the idea that plants are unique among living things because of their ability to produce their own food and that – because of this – the food that plants produce is the ultimate source for all living things on Earth. In addition, at this level, students begin to link the gas exchange which is useful to people to the plant's own processes. Finally, Level 6 represents full understanding of the *Plants as Producers* progress variable. In addition to the understandings at Level 5, students at this level can identify the type of food that plants produce (glucose, or sugar). They recognize that the process of photosynthesis allows plants to use light energy to create food and, thus, the critical role of the Sun for all life on Earth. They also understand that the process of photosynthesis uses carbon dioxide (which people breathe out) and produces oxygen (that people breathe in). In addition, students at this level, recognize that plants also require oxygen and release carbon dioxide; respiration is not included in this learning progression, so students are not expected to understand why.

Conclusions and Implications

This paper contributes a specific learning progression in an area of the curriculum in which relatively little learning progressions work has been conducted (early elementary school) and for a topic of great importance in the elementary school curriculum. As discussed briefly above, this is only the first step in a much longer process of revision and validation. The interview protocol will need to be modified in light of the revised learning progression, and additional interviews (with a wider range of students) will need to be conducted to determine whether it adequately captures students' ideas. In addition, because cognitive interviews can feasibly be conducted with only a limited number of students, paper-and-pencil assessment items have been developed and administered using a "think aloud" protocol (Ericsson & Simon, 1993). Analysis of students' responses to these items will be used to inform both the learning progression and the associated assessment items. Multiple steps of a cycle involving the learning progression, cognitive interviews, and administration of paper-and-pencil items (including some "think aloud" interviews) will be required to be confident that the learning progression captures student thinking. Additional work, including longitudinal data – at a minimum, following students' development in a single curriculum unit on plants – would be required to verify that students' ideas actually progress in hypothesized ways.

Part of this revision process must include an issue which – although considered in the revision process reported here – has not been fully explored in this paper: the extent to which students respond consistently at a given level of the learning progression. This issue actually consists of two related concerns. First, each of the progress variables describes student thinking about a set of related ideas. (For example, *Plants as Producers* includes both ideas about plants' independence and about multiple ways in which people depend upon plants.) However, there is little empirical evidence to indicate how these ideas actually "hang together" in students' minds. We have hypothesized, for example, that students who recognize that – under the right conditions – plants can grow independent from other living things will also recognize that people eat plants for food and that plants help people to breathe. Further data is required to test this hypothesis, and revisions to the learning progression must take this data into consideration. In addition, there is a loose assumption that students who hold ideas at Level 2 of the *Plants as Producers* progress variable, will also hold ideas at Level 2 of the *Food for Plants* and *Energy for Plants* progress variables. Again, this is a hypothesis which must be tested empirically. Second, particularly in physics (e.g., Halloun & Hestenes, 1985), there is overwhelming evidence that students do not apply ideas consistently across contexts. Thus, during the course of a single interview or set of paper-and-pencil items, students may not consistently express understandings at the same level of the learning progression. This is a challenge for all learning progressions work; if students do not respond consistently, then a reliable diagnosis of their learning progression level is not possible. While this problem may persist even when levels are defined to be as consistent with student thinking as possible, it is important for the revision process to carefully consider student responses which seem at odds with other indicators of their learning progression level.

Eventually, a well-validated learning progression may have important classroom implications. We are particularly concerned by a pattern observed in our data. Students reported learning about plants every year; however, the same naïve ideas about plants were repeated throughout our sample. It did not appear that students were deepening their understanding of plant nutrition over the course of their elementary school science experiences. Indeed, those students with unusually high levels of understanding of plant nutrition tended to report experiences outside of school which contributed to that understanding. Thus, it is clear that current instruction – at least in the school participating in this study, but likely in other schools as well – is failing to support students' development of solid conceptual understandings about plant nutrition. The Plant Nutrition learning progression may be used to inform instruction – including formative assessment practices during instruction – which may change these patterns. Making teachers aware of the steps that students may take in reaching full understanding of this topic provides both intermediate targets for instruction and typical ideas to listen for (and engage with) during instruction. Our ultimate goal is for this learning progression to be used by teachers such that students make progress in their understanding of plant nutrition in elementary school and are well-positioned to understand more advanced topics, such as carbon transformation and ecological systems. These understandings are of crucial importance for environmental literacy and, thus, for informed decision-making about environmental issues.

References

Alonzo, A. C. (2009, April). Design criteria for learning progressions to support teachers' formative assessment practices. In A. C. Alonzo (Chair), *Learning progressions in science: Tools for teaching, learning, and assessment*. Symposium conducted at the annual meeting of the American Educational Research Association, San Diego, CA.

Alonzo, A. C., & Steedle, J. T. (2009). Developing and assessing a force and motion learning progression. *Science Education, 93*, 389-421.

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. New York: Oxford University Press.

Anderson, C. W., Sheldon, T. H., & Dubay, J. (1990). The effects of instruction on college nonmajors' conceptions of respiration and photosynthesis. *Journal of Research in Science teaching, 27*, 761-776.

Barker, M. & Carr, M. (1989). Teaching and learning about photosynthesis. Part 1: An assessment in terms of students' prior knowledge. *International Journal of Science Education*, 11(1): 49-56.

Barman, C. R., Stein, M., McNair, S., & Barman, N. S. (2006). Students' ideas about plants and plant growth. *The American Biology Teacher, 68*(2), 73-79.

Bell, B. (1985). Students' ideas about plant nutrition: What are they? *Journal of Biological Education, 19*, 213-218.

Black, P., & Wilson, M. (2007, April). The idea of a learning progression as a core for both instruction and assessment. In A. C. Alonzo (Chair), *Assessing understanding through the use of learning progressions*. Symposium conducted at the annual meeting of the American Educational Research Association, Chicago.

Briggs, D. C., Alonzo, A. C., Schwab, C., & Wilson, M. (2006). Diagnostic assessment with ordered multiple-choice items. *Educational Assessment, 11,* 33-63.

Cañal, P. (1999). Photosynthesis and 'inverse respiration' in plants: An inevitable misconception? *International Journal of Science Education, 21*, 363-371.

Çepni, S., Taş, E., & Köse, S. (2006). The effects of computer-assisted material on students' cognitive levels, misconceptions and attitudes towards science. *Computers & Education, 46*, 192-205.

Ericsson, K.A., & Simon, H.A. (1993). *Protocol analysis: Verbal reports as data.* Cambridge, MA: The MIT Press.

Furtak, E. M. (2009, April). Learning progressions to support teacher development. In A. C. Alonzo (Chair), *Learning progressions in science: Tools for teaching, learning, and assessment*. Symposium conducted at the annual meeting of the American Educational Research Association, San Diego, CA.

Gotwals, A. W., & Songer, N. B. (2008, March). Scenario-based assessment: Gathering evidence of student reasoning in biodiversity. In A. C. Alonzo & A. W. Gotwals (Chairs), *Diverse perspectives on the development, assessment, and validation of learning progressions in science.* Symposium conducted at the annual meeting of the American Educational Research Association, New York.

Halloun, I.A., & Hestenes, D. (1985). Common sense concepts about motion. *American Journal of Physics, 53*, 1056-1065.

Leach, J., Driver, R., Scott, P., & Wood-Robinson, C. (1996). Children's ideas about ecology 2: Ideas found in children aged 5-16 about the cycling of matter. *International Journal of Science Education, 18*, 19-34.

LessonLab Research Institute. (2008, August). *Videocases for Science Teaching Analysis (ViSTA): Plants*. Santa Monica, CA: Author.

Marmaroti, P., & Galanpoulou, D. (2006). Pupils' understanding of photosynthesis: A questionnaire for the simultaneous assessment of all aspects. *International Journal of Science Education, 28*, 393-403.

Mohan, L., Chen, J., Anderson, C. W. (2009). Developing a multi-year learning progression for carbon cycling in socio-ecological systems. *Journal of Research in Science Teaching, 46*, 675-698.

National Assessment Governing Board. (2006). *Science framework for the 2009 National Assessment of Educational Progress* (Prepublication Edition). Retrieved July 11, 2006 from http://www.nagb.org/pubs/naep_science_framework_2009.doc

National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.

National Research Council. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: The National Academies Press.

National Research Council. (2006). *Systems for state science assessment*. Washington, DC: The National Academies Press.

National Research Council (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: The National Academies Press.

National Science Foundation. (2005). *Instructional Materials Development program solicitation* (NSF 05-612). Arlington, VA: Author. Retrieved August 19, 2009, from http://www.nsf.gov/pubs/2005/nsf05612/nsf05612.htm

National Science Foundation. (2006). *Discovery Research K-12 program solicitation* (NSF 06-593. Arlington, VA: Author. Retrieved August 19, 2009, from http://www.nsf.gov/pubs/2006/nsf06593/nsf06593.pdf

National Science Foundation. (2009). *Research and Evaluation on Education in Science and Engineering program solicitation* (NSF 09-601). Arlington, VA: Author. Retrieved August 19, 2009 from http://www.nsf.gov/pubs/2009/nsf09601/nsf09601.txt

Özay, E., & Öztas, H. (2003) Secondary students' interpretations of photosynthesis and plant nutrition. *Journal of Biological Education, 37*(2), 68-70.

Roberts, L., Wilson, M., & Draney, K. (1997, June). *The SEPUP assessment system: An overview*. BEAR Report Series, SA-97-1. University of California, Berkeley.

Schwarz, C., Reiser, B., Davis, E., Kenyon, L., Acher, A., Fortus, D., et al. (2009). Developing a learning progression for scientific modeling: making scientific modeling accessible and meaningful for learners. *Journal for Research in Science Teaching*, 46*,* 632-654.

Shavelson, R. J., Yin, Y., Furtak, E. M., Ruiz-Primo, M. A., Ayala, C. C., Young, D. B., et al. (2008). On the role and impact of formative assessment on science inquiry teaching and learning. In J.E. Coffey, R. Douglas, & C. Sterns (Eds.), *Assessing science learning: Perspectives from research and practice* (pp. 21-36). Arlington, VA: NSTA Press.

Smith, C. L., Wiser, M., Anderson, C. W., & Krajcik, J. (2006). Implications of research on children's learning for standards and assessment: A proposed learning progression for matter and the atomic molecular theory. *Measurement: Interdisciplinary Research and Perspectives, 14*, 1-98.

Smith, E. L., & Anderson, C. W. (1984). Plants as producers: A case study of elementary science teaching. *Journal of Research in Science Teaching, 21*, 685-698.

Songer, N., Kelcey, B., & Gotwals, A. (2009). How and when does complex reasoning occur? Empirically driven development of a learning progression focused on complex reasoning about biodiversity. *Journal of Research in Science Teaching*, 46, 610-631.

Stevens, S.Y., Shin, N., Delgado, C., Krajcik, J., & Pellegrino, J. (2007, April). Developing a learning progression for the nature of matter as it relates to nanoscience. In A. C. Alonzo (Chair), *Assessing understanding through the use of learning progressions*. Symposium conducted at the annual meeting of the American Educational Research Association, Chicago. Retrieved May 21, 2008, from http://hi-ce.org/presentations/documents/UM_LP_AERA_2007.pdf

Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Mahwah, NJ: Lawrence Erlbaum Associates.

DEFINING LEVELS OF SCIENTIFIC INQUIRY SKILLS IN LOWER SECONDARY BIOLOGY EDUCATION

Andrea Möller
Justus-Liebig University Giessen

Jürgen Mayer
University of Kassel

Abstract

The detailed developmental processes of scientific inquiry competence within the school environment have hitherto been studied only unsystematically. In the here presented study, we predict and measure five qualitative levels of scientific inquiry competence in the four central skills "formulating questions", "generating hypotheses", "investigation planning", and "interpreting data" based on a differentiation through levels of complexity and a qualitative grading according to problem-solving processes. On the basis of 24 open test items a multi-matrix design was used to perform a longitudinal test on the inquiry competence of 1129 German students (age 10-16). We found all five predicted qualitative competence levels present in the students' test answers. Detailed analyses reveal that students' performance development differs significantly within the investigated grades and within the four skills. The expected increase of inquiry competence within one school year is due to a significant qualitative increase within the skill levels. The results support a qualitative grading of each of the four skills according to scientific problem-solving processes and provide more precise information about the qualitative increase within the skills. The described threshold values for each competence level permit and facilitate a targeted choice of test items. Our study can thus help providing an assessment tool for individual promotion of scientific inquiry skills.

Introduction

Providing students with competences of scientific inquiry represents a key goal of science education in general and of biology education in particular. However, results of international student assessment programs like TIMSS or PISA as well as several research studies in the field of science education show that German students in lower secondary biology education (grades 5-10) still lack competence especially in the field of scientific problem solving and science process skills (e.g. Mayer et al. 2003, Hammann 2006). As one reaction to these studies, inquiry competence was being officially integrated in the 2004 German Federal Education Standards (KMK 2004). It can, nevertheless, still be shown that scientific problem solving, especially planning and conducting experiments, is either widely neglected or taught unsystematically in German schools (e.g. Mayer 2004). Therefore it is not surprising that recent studies show that students of all ages still lack crucial competences in the process of scientific inquiry. They plan their experiments unsystematically, do not control the variables and hardly ever include experimental controls (Hammann et al. 2008). Over 50% do not meet the requirements in all four central skills "formulating questions", "generating hypotheses", "planning an investigation", and "interpreting data" (Möller et al. 2008, 2009). In order to address the students' problems and create a learning environment that encourages the individual to participate in the science process, it appears crucial to focus on finding out how inquiry competences can be described and assessed in greater detail.

Rationale

Scientific inquiry is mainly described as a problem-solving process that involves several relevant skills (e.g. Klahr 2000, Gott & Duggan 2001, Abd-El-Khalik et al. 2004, Mayer 2007). Möller, Grube & Mayer (2008) were able to empirically support the assumption that scientific inquiry can be differentiated in four central skills: "formulating questions", "generating hypotheses", "investigation planning", and "interpreting data". However, the exact developmental processes of scientific inquiry competence within the school environment remain subject to further study. Particularly, the analysis of qualitative levels within the mentioned four central skills awaits detailed study. In order to establish a grading of competence levels within inquiry skills, several criterions are suggested. For example, Bybee (2002) proposes to use a gradual shift of students' concepts of everyday life to scientific concepts. Others focus on the students' use of relevant variables (Schauble et al. 1992) and the more or less systematic approach during investigation (Hammann 2004). In the here presented study, we predict and measure five qualitative levels of inquiry competence which are based on a differentiation through levels of complexity and a qualitative grading of each skill according to problem-solving processes (Möller et al. 2007, Mayer et al. 2008). Longitudinal studies over all grades in lower secondary biology education show an overall significant increase of inquiry competence (Möller et al. 2009). However, it is not clear to which extend this is due to a qualitative increase within the skill levels.

Therefore our main research questions are: 1) Can all five predicted qualitative competence levels for each of the four inquiry skills be found in the students' test answers? and 2) Is the expected increase of inquiry competence within one school year due to a *qualitative* increase within the skill levels?

Methods

The study employed a longitudinal test design with two test points over one school year in order to test the levels of inquiry competence of students from grade 5-10 (age 10-16). A paper-and-pencil test was conducted with 1129 students at the start and end of the school term, respectively. Both tests were conducted in a multi matrix design with 24 open test items, equally representing the four inquiry subskills "formulating questions," "generating hypotheses," "planning an investigation," and "interpreting data," which were already confirmed in previous studies (Grube et al. 2007). Each individual test included six items. Data were analyzed with the help of specific coding schemes for each of the four skills (intercoder reliability: *Cohens Cappa* = 0.88 - 1.0 for all proposed skills). Up to 14 individual codes (ranked in degree of difficulty) were combined to five assumed levels of competence according to increased level of difficulty and problem-solving processes. Data from both tests were entered and analyzed using *ConQuest* (Wu, Adams & Wilson 1997) for item-response modelling and drawing plausible values. Plausible values represent a student's inquiry competence by integrating personal ability and item difficulty. In order to achieve threshold values for each skill level we defined the threshold at 50%, meaning that a student must be able to solve at least half of the items within one skill in order to achieve the next level. This method is used in accordance to the definition of threshold values in international large-scale student assessment studies, such like TIMSS and PISA. Further analyses were conducted using the Statistical Software Program Package (*SPSS*) 15.0.

Results

At the beginning of the school term 45.1% of all students across all grades achieve the first competence level, while 11.8 % are below level 1, meaning that they do not reach any level of scientific inquiry competence. 37.7 % reach the second, and 5.4 the third level of competence, the latter representing overall the requirements of the national and international science education standards. No student reaches level 4 and 5. At the end of the term, the percentage of students that are below level 1 decreases significantly to 8.1 ($p < .001$, Wilcoxon). Also, the number of students that reach level 1 decrease significantly from 45.1 to 31.2 % ($p < .001$, Wilcoxon). At the same time the percentage of students that reach level 2 and 3 increase significantly over all four inquiry skills (8 and 8.8 %, respectively; both $p < .001$, Wilcoxon) (see fig.1). At the second test 0.8 % of the students reach level 4. Although some students reached level 5, numbers were low and highly inconsistent in distribution among grades. So far, it cannot be included in the *item-response* model. In all four inquiry skills and throughout all age groups, less than 50% of all students achieved level 3.

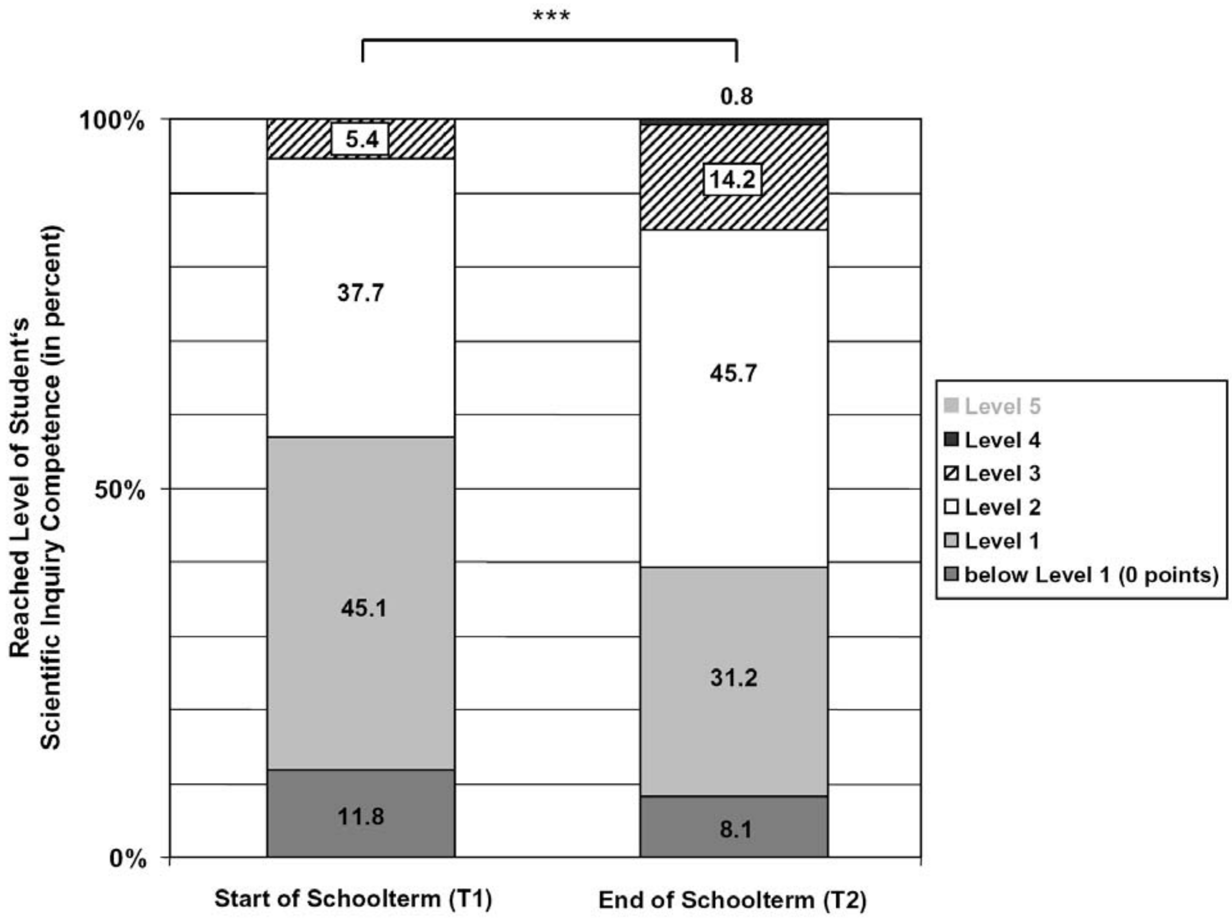


Fig. 1: Reached levels of students' scientific inquiry competence over all inquiry skills and grades 5 -10 (N = 1129) at the beginning (T1) and end of the schoolterm (T2).

Conclusions and Implications

First, the five predicted qualitative competence levels for each of the four inquiry skills were all present in the students' test answers. Therefore, the results of this study support the predicted qualitative grading of each of the four skills according to scientific problem-solving processes. The open test items used in this test can be applied on all accounts to evaluate inquiry competence up to level 3 that represents the requirements of national education standards. However, competence levels 4 and 5 could not be included in the model. We believe that those two highest competence levels cannot be evaluated by paper pencil tests alone but probably need other assessment methods such like interviews or video analysis.

Second, the expected increase of inquiry competence within one school year is due to a significant *qualitative* increase within the skill levels (p <. 001 over all levels). Therefore, in addition to a prediction about the overall competence increase of each student, our model allows more precise information about the qualitative increase within the skills. For science teachers, an accurate assessment of students' scientific inquiry skills is crucial in order to gain insight in the progress of each individuals' performance. It facilitates lesson planning and enables the teacher to provide accordant feedback, as well as distribute adequate exercise units. With a greater knowledge of students' difficulties and needs, novel teaching concepts can be developed that are possibly adjusted to different performance levels of scientific inquiry and still meet the actual classroom setting.

Third, the defined threshold values for each competence level permit and facilitate a more targeted choice of test items in future assessments on students' scientific inquiry competence.

References

Abd-El-Khalik, F. et al. (2004). Inquiry in Science Education: International Perspectives. *Sci Ed 88*:397-419.

Bybee, R.W. (2002). Scientific Literacy – Mythos oder Realität? In: W. Gräber, P. Nentwig, T. Koballa, R. Evans (eds), Scientific Literacy. Der Beitrag der Naturwissenschaften zur Allgemeinen Bildung. Opladen: 2002.

Gott, R. & Duggan, S. (2001). Understanding Scientific Evidence – Why It Matters and How It Can Be Taught. In: Ratcliffe, M. (1998): ASE Guide Secondary Science Education, Hatfield 2001.

Grube C., Möller A., Mayer J. (2007): Dimension eines Kompetenzstrukturmodells zum Experimentieren. In: Bayrhuber et. al. (Hrsg.): Ausbildung und Professionalisierung von Lehrkräften. Univ. Kassel, Kassel: S. 31-34.

Hammann, M. (2004). Kompetenzentwicklungsmodelle: Merkmale und ihre Bedeutung - dargestellt anhand von Kompetenzen beim Experimentieren. *MNU 57*/4:196-203.

Hammann, M.; Phan, T.T.H.; Ehmer, M. & Bayrhuber, H.(2006). Fehlerfrei Experimentieren. *MNU 59*/5:292-299.

KMK (2004). Bildungsstandards im Fach Biologie für den mittleren Schulabschluss. München: Wolters Kluwer.

Klahr, D. (2000). Exploring Science. The Cognition and Development of Discovery Processes. Cambridge: MIT Press.

Mayer, J., Keiner, K. & Ziemek, H.-P. (2003). Naturwissenschaftliche Problemlösekompetenz im Biologieunterricht. In: Bauer, A. et. al. (eds.), Entwicklung von Wissen und Kompetenzen im Biologieunterricht. Internationale Tagung der Sektion Fachdidaktik im VdBiol. Berlin 2003, Kiel: IPN, pp 21-24.

Mayer, J. (2004). Qualitätsentwicklung im Biologieunterricht. MNU 57/2:92-99.

Mayer, J. (2007). Erkenntnisgewinnung als wissenschaftliches Problemlösen. In: Krüger, D. & Vogt, H.: Handbuch der Theorien in der biologiedidaktischen Forschung. Berlin/Heidelberg: Springer, pp 179-188.

Mayer, J., Grube, C., & Möller, A. (2008). Kompetenzmodell naturwissenschaftlicher Erkenntnisgewinnung. In: U. Harms & A. Sandmann (Eds.), Lehr- und Lehrnforschung in der Biologiedidaktik – Ausbildung und Professionalisierung von Lehrkräften. Innsbruck: Studienverlag, pp. 63-79.

Möller, A., Grube, C., Mayer, J. (2007). Kompetenzniveaus der Erkenntnisgewinnung bei Schülerinnen und Schülern der Sekundarstufe I In: Bayrhuber et. al. (Hrsg.): Ausbildung und Professionalisierung von Lehrkräften. Kassel: Univ. Kassel, pp. 55-58.

Möller, A., Grube, C. & Mayer, J. (2008). Skills and Levels of Students' Inquiry Competence in Lower Secondary Biology Education (grade 5-10). 2008 International Conference of the National Association of Research in Science Teaching (NARST) "Impact of Science Educational Research on Public Policy". Baltimore, MD.

Möller A., Grube C., Hartmann S. & Mayer J. (2009): Increase of inquiry competence: a longitudinal large scale assessment of students' performance from grade 5 to 10. 2009 International Conference of the National Association of Research in Science Teaching (NARST) "Grand Challenges and Great Opportunities in Science Education", Garden Grove, CA.

Schauble, L. et al. (1992). The Integration of Knowledge and Experimentation Strategies in Understanding a Physical System. Applied Cognitive Psychology 6: pp 321-343.

NRC (1996). National Science Education Standards. National Academy Press, Washington, DC.

Wu, M. L., Adams, R. J. & Wilson, M. (1997). ConQuest - Generalised item response modelling software, Draft Release 2. Camberwell: Australian Council for Educational Research.

ALTERNATIVE THINKING STYLES OF STUDENTS GRADUATED FROM PRIMARY SCHOOL IN THE CONCEPT OF ENERGY

Hasan Özcan
Aksaray University

Mustafa Sabri Kocakülah
Balıkesir University

Abstract

This study aims to reveal the alternative thinking styles of students, who finished up their eight year long primary education, about the subjects of energy. For this aim, a conceptual understanding test was developed by the researchers with the help of a comprehensive analysis on literature. This conceptual understanding test, which was developed in order to reveal the students' who constitute the sample alternative thinking styles about energy transformation which occurs during the transition of a material, the conservation of a energy and the relationship between the chemical bond and energy, consist of 4 questions in two stages. The data gathered from the research was appraised via qualitative analysis

Introduction

Preparing students for the life and for the higher education by developing their interest and skills is among the main aims of primary school education. This aim also requires that students abstain from the negative effects of the errors which become permanent as a result of bringing the problems met in one education stage to the higher education stage and the learned things' becoming old, with an attribution to the learning by living and doing principle. The concept of energy, which exists not only in primary education but also in the life an individual before and after the primary education, and its subjects related with education are among the keywords which one always comes across. When its role concerning with is taken into consideration, the concept of energy takes place among the concepts which sum and reflect the teaching and learning of science. Nevertheless when it is looked from a broader perspective, the learning of energy begins with being acquainted with some concepts before education just as it is so in science (Watts, 1983). With have been the subject of recent studies; in other words, to the analysis of individual alternative thinking process. There are many studies which comply with the subjects of energy in this regard and are conducted with identical cultures and results (Watts, 1983; Lijnse, 1990; Maskill, Pedrosa & Maskill, 1997; Liu & Tang, 2004).

Literature Review

In a study conducted by Maskill and Pedrosa (1997) which included 183 students from 6 classes who got education in Portuguese and whose ages ranked between 15-16, there were found student discourses about energy like "it transits", "it disappears" and "it is kept in the body". Related with these findings, it was revealed that students made explanations with daily language disregarding the conservation of energy. Researchers note that the difficulty in the instruction of abstract and complex concepts stems from previous and inadequate knowledge and the resistance of daily language towards scientific language. A study by Boyes and Stanisstreet (1991) also backs up this resistance. In this study, the ones who take science course and who do not in the daily life and to which extent the takers can differentiate scientific language and daily language are put forward. According to the results gained from this study, while some of the students are aware of that plants take energy from the sun, % 31 of the sample declared that plants take energy from the ground, % 28 argued that they take energy from water and % 28 declared that they take energy from the atmosphere (Boyes & Stanisstreet, 1991).

The close relationship between the concept of energy and life also brings forward some unfavorable results as well as favorable ones. The thinking structures which form the basis of the problems that may occur are listed by Watts and Gilbert (1982) in seven alternative frameworks (Watts, 1983).

- Human-centered framework. It includes the belief concerned with human and especially refers to being energetic like "having a plenty or losing energy so getting tried"

- Storage framework. While some materials have energy and they can be recharged, some others need energy and they consume the energy they get.

- Energy is a material-component or constituent. Energy is an inanimate component that requires to be triggered in materials and incidents just as in the incidence of combustion. For instance energy exists in food but it becomes operative only after you eat it.

- Activity framework. It is the reflection of energy towards the exterior of the activity especially during movement.

- Production framework. Energy is neither a component nor a process. However; it is the by-product of a production or a disappearance.

- Functional framework. Energy is regarded as a general type of a fuel which facilities the life.

- Course transference framework. Energy is regarded as a type of liquid which is transferred in some processes.

Aim of The Study

This research aims to identify the alternative thinking styles of students, who graduated from primary school, concerning the concept of the energy and its related subjects. Research especially emphasizes on defining the students' alternative thinking styles about, the conservation of a energy, the energy transformations which occur in the material at macro and micro extents during transition, the course of energy in the ecosystem and the relationship between the chemical bond and energy. The alternative beliefs are presented as data, based on the sentences inherent in the explanations of students.

Method

210 graduate students from 5 different primary schools constituted the sample of this research. The conceptual comprehension test, which was used in the study, consisted of two open-ended questions taken from the studies carried out by Duit (1984) and Swackhamer, Hestenes (2005) and 4 questions, 3 of which were open-ended and 1 was a multiple-choice question, developed by the researchers. This research, which is dominantly qualitative, has a descriptive analysis pattern (Duit, 1984; Swackhamer & Hestenes, 2005). In the analysis process of the results gained from the research, the context analysis was followed and each student was assigned with a discrete code number from the numbers of the sample interval.

The Findings and Comments

The findings of the research are presented with four different titles in tables. The first of them which is categorized in terms of energy transformation which includes transformations which includes main subject titles such as energy types, energy transmission and conservation of energy is presented in Table 1. As the confusion of the concept of energy with other concepts is given under one item (Item 14) No other category concerned with this concept has been created and it is investigated in Table 1.

Table 1 Defined Alternative Thinking Styles About The Matter of Energy Transformation

1	Not knowing the functions of the energy types.
2	Confusing the energy types.
3	Regarding energy transformation as the decrease and increase in energy.
4	Explaining an energy transformation incidence with daily language rather than scientific language as a result of not comprehending the energy transformation fully.
5	Challenging in comprehension of a energy transmission and the supersession of scientific language by daily language.
6	The perception that there is no energy in a frictionless environment as there is no electricity.
7	The perception that energy never changes in frictionless environments as it changes just through friction.
8	The belief that the total energy of a material which starts moving from any point certain speed in a frictionless environment reduces.
9	The belief that the total energy of a material which starts moving from any point certain speed in a frictionless environment, increases.
10	The belief that the total energy of a material which starts moving from any point certain speed in a frictionless environment, changes according to the shape, elevation and height it moves on.
11	The belief that the total energy of a material changes just with the duration of movement.
12	The belief that if the speed of two materials is equivalent, their energy is always equivalent, too.
13	The belief that energy increases in accordance with the increase in weight.
14	Confusing concept of energy with the concept of power.

According to the Table 1, it is found out that students, constituting the sample, face difficulty in the comprehension of the functions of the energy types and some students confuse these energy types. There are many delusions of students in the literature that accompany with this study (Kruger, 1990; Trumper, Raviolo & Shnersch, 2000; Stylianidou, Ormerod & Ogborn, 2002; Trumper, 1997; Kocakülah, Üstünlüoğlu & Kocakülah 2005). For instance whereas it found that the mechanic energy just equals with kinetic energy valuation and the comparison of kinetic energy and potential energy was detected in this study, the literature emphasizes on that its not known that mechanic energy is the some as the total of kinetic and potential energy and the delusion that kinetic energy does not depend on speed (Kruger, 1990; Gülçiçek & Yağbasan, 2004). The alternative thinking, which is presented in Table 1 and coincides with item 3, accompanies with the delusion put forward by Hapkiewicz (1992) that energy decrease in the energy transformation (Vargo, 1997). Moreover, the findings, which have been found out by this study and is not present in literature and which suggests that energy increases in the energy transformation, contributes to the literature via this study. Nevertheless, the usage of daily language in the explanation of energy transformation rather than scientific language (Item 4-5) takes a considerable place in literature (Duit, 1984; Kocakülah, Üstünlüoğlu & Kocakülah 2005).

In addition, the delusion concerned with the relationship between friction energy found in the items 6., 7., 8., 9., and 10. in the Table 1 suggests that there is a concept confusion stemming from the meaning of frictionless environment and belief that energy stays stable in frictionless environment, and the belief this concept confusion occurs between the concepts friction and electricity. It is among the findings of this study accompanying with the literature that the principle of energy conservation is not known with all of its components and the delusion that energy changes just through speed or weight leads to a confusion between the concept and concept of power (Kruger, 1990; Kocakülah, Üstünlüoğlu & Kocakülah 2005; Vargo, 1997; Küçük, Çepni & Gökdere, 2005; Kruger, Palacio & Summers, 1992; Summers, Kruger, Mant & Childs, 1998; Griffiths &Preston, 1989).The findings of the study are secondly examined under the title of alternative beliefs about the changes that occur in the physical structure of materials.

In Table 2, it is seen that the students, forming the sample concerned with the evaluation of changes that occur in the physical structure of a material in the dimension of a energy, have 7 alternative beliefs. These beliefs can be listed via a macro perspective (Item 1, 2 and 7)

Table 2 Defined Alternative Thinking Styles About The Changes That Occur In The Physical Structure of A Materials

1	The belief that as the concrete form of a material is more orderly than its liquid form and the liquid form of it is more orderly than its gas form, the energy that its granules have is much more.
2	The belief that as the concrete form of a material is more intense and harder than its liquid form and its liquid form is more intense and harder than its gas form; the energy than its granules have is much more.
3	The granules of a material can not move when it is in concrete form. When the material is in the form of liquid, the granules can move a bit and when it is in the form of gas, the granules are more active. The belief because of this reason, the energy of granules in concrete form is much more with respect to the gas form.
4	The distance between the granules is longer in the gas form of a material than is in its liquid form and is longer in the liquid form than in its concrete form. The belief that because of this reason, energy of granules in concrete form is much more than its liquid form and the energy of granules in liquid form is much more than its gas form.
5	As the attraction power of the granules of a concrete material is much more than the attraction power of the granules of the liquid form of the same material, the energy that the granules of the concrete material have is much more. The belief that although the attraction power of the granules of the liquid form of a material is less than the concrete form, as it is more than the gas form in which there is almost no attraction power between the granules, the energy it has is much more than the gas form.
6	As the attraction power of the granules and the distance between the granules are small in the concrete form of a material with respect to liquid form and gas form, the energy its granules and the distance between them are big, the energy the granules of a material have in gas form is much more.
7	As the concrete, liquid and gas forms of a material are the different forms of the same material, the energy the granules have in all of these forms remains stable.

The difficulties resulting from such approaches toward the matter of transition are also found in literature (Griffiths &Preston, 1989; Ben-Zvi, Eylon & Silbestein, 1986; Brook & Drıver, 1984; Novick & Nusbaum, 1981;

Shepherd & Renner, 1982; Kruger & Summers, 1989). Another point in which the study results accompany with literature is the belief that the granules of concrete materials are inactive, hard, intense and indecisive and there is a little or no space between the granules (Lee, Eıchınger, Anderson, Berkhemer & Blakeslee, 1993; Stavy, 1991). From this point of view, the belief that students declare wrongfully that the energy the granules of its liquid form have is much more than energy of the granules of the gas form have is not found in literature.

On the other hand, another matter, in which students constituting the sample face difficulty in terms of the concept of energy, is the break and formation of chemical bond Table 3).

Table 3 Defined Alternative Thinking Styles in The Matter of The Relationship Chemical Bond and Energy

1	The bond in the H_2 molecule (H-H) breaks during the formation of water. The belief that energy emerges in this process.
2	There appears a bond between the H_2 molecules (H-H) during the formation of water. The belief that energy is required in this process.
3	There appears a bond between the H_2 mol in the H_2 molecule (H-H) breaks during the formation of water. The belief that energy emerges in this process.
4	The bonds in the H_2O molecule (H-O and O-H) are formed during the formation of water. The belief that energy emerges in this process.
5	The bonds in the H_2O molecule (H-O and O-H) break during the formation of water. The belief that energy emerges in this process.
6	The bonds in the H_2O molecule (H-O and O-H) break during the formation of water. The belief that energy is required in this process.
7	When the bonds between the H_2O molecule (H-O and O-H) are created during the formation of water, energy is not required, the process occurs by itself.

Nevertheless, the alternative beliefs concerned with the fact that in the break oh chemical bonds, which also take an important place in literature, energy emerges and energy is required in the formation of them are also handled in this study (Item 1 and 4) (Gilbert, Osborne, & Fensham, 1982; Ross, 1993; Barker, 1995; De Vos & Verdonk, 1986; Hong Kwen Boo, 1998). Additionally, the alternative beliefs stated in the items 2., 3., 5., 6. and 7. in Table 3 contributed to the present literature.

In addition, the alternative beliefs of students graduated from primary school about the explanation of energy course in the food-energy pyramid are summarized in Table 4 with the existent published literature about this matter, this study expands the findings of literature in some points (Lin & Hu 2003; Anderson, Sheldon & Dubay, 1990; Özay & Öztaş, 2003). According to this knowledge, the alternative beliefs about the explanation of the energy course in the food-energy pyramid in terms of "the diet of living beings", which constitutes the item 1 in Table 4 and is detected by students graduated from primary school and university, and the alternative beliefs about the wrong explanation of this course in terms of "the conservation of energy" and "the valuation of food" are supported by literature (Adeniyi, 1985; "stevetrash", 2006). Moreover, the students' beliefs about the explanations given about the energy course in the food-energy pyramid in terms of the dimensions seen in Table 4 such as "the physical size of the living beings", "the part of the pyramid which belongs to the livings beings" and "the capacity of consuming of the living beings" take place in the literature accompanying with the explanation, "while going to superior organisms from primitive organisms, the amount of energy that is had increase" (Marek, 1986) Also the alternative belief, which is presented as item 4 in table 4 and which the explanation of energy discourse in food-energy pyramid is made on the basis of "the number of the living beings" by students graduated from both primary school and university, complies with the findings of the literature (Munson, 1994).

Table 4 Defined Alternative Thinking Styles in The Matter of Energy Course in The Ecosytem

	The energy course which occurs in the food-energy pyramid;
1	Its explanation on the basis of "the diet of the living beings".
2	Its explanation in terms of "the physical sizes of the living beings".
3	Its explanation in terms of "the part of the pyramid which belongs to the living beings".
4	Its explanation on the basis of "the numbers of living beings".
5	It Its explanation on the basis of "the capacity of consuming of the living beings".
6	Its explanation in terms of "the moving skills of the living beings".
7	Its explanation in terms of "the conservation of energy".
8	It Its explanation on the basis of "the valuation of food".

Results and Suggestions

The alternative beliefs of students graduated from primary school about energy transformation, energy changes which occur in macro and micro dimensions in a material during the transition, the conservation of energy, the energy course in the ecosystem and the relationship between chemical bond and energy are presented in Table 4 in the framework of the findings gained from the research. While some of these alternative beliefs put forward are supported by literature some are peculiar to this research and with its this aspect, it brings forward new approaches to the literature.

Teachers are assigned with important roles in overcoming the learning difficulties about the subjects of energy and in the correction of these alternative beliefs which are created at the end of this process. In this aspect, it is crucial to analyze the literature focusing research carried out in Turkey and abroad elaborately prior to the instruction, to detect the concepts that students have before instruction and to take necessary precautions about the scientifically wrong beliefs, to give more importance to the new techniques (plays, scientific stories, etc.) which will provide a more enjoyable classroom environment with a permanent learning. Moreover, the interdisciplinary character of the energy concepts should be taken into consideration by curriculum developers in order to provide consistency between these concepts at the levels of pre-school, primary school and secondary school.

References

Adeniyi, B. O. (1985) Misconceptions of selected ecological concepts held by some Nigerian students. *Journal of Biological Education*, 19(4), 311-316.

Anderson, C.W., Sheldon, T.H., & Dubay, J. (1990). The effects of instruction on college nonmajors' conceptions of respiration and photosynthesis. *Journal of Research in Science Teaching*, 27(8), 761-776.

Barker, V. (1995). A longitudinal study of 16-18 year olds' understanding of basic chemical ideas. Unpublished M.Phil. Thesis, Department Of Educational Studies, University of York.

Ben-Zvi, R., Eylon, B., & Silbestein, J. (1986). Is an atom of copper malleable? *Journal of Chemical Education*, 63(1), 64-66.

Boyes, E., & Stanisstreet, M. (1991). Misconceptions in first-year undergraduate science students about energy sources for living organism. *Journal of Biological Education*, 25(3), 20-213.

Brook, A. B., & Driver, R. (1984). Aspects of secondary students understanding of the particulate nature of matter. Leeds: University of Ledds.

De Vos, W., & Verdonk, A. (1986). A new road to reactions, Part 3: Teaching the heat effect of reactions. *Journal of Research in Chemical Education*, 63, 972–974.

Duit, R. (1984). Learning the energy concept in school–emprical results from the Philippines and West Germany. *Physics Education*, (19), 59-66.

Gilbert, J. K., Osborne, J., & Fensham, P. J. (1982). Children's science and its consequences for teaching. *Science Education*, 66(4), 623-633.

Griffiths, A. K., & Preston, K. R. (1989). Models of molecules and atoms. Paper presented at National Association for Research in Science Teaching.

Gülçiçek, Ç., & Yağbasan, R. (2004). Basit sarkaç sisteminde mekanik enerjinin korunumu konusunda öğrencilerin kavram yanılgıları. *Gazi Eğitim Fakültesi Dergisi*, 24(3), 23-38.

Hong Kwen Boo, H. (1998). Students' understandings of chemical bonds and the energetics of chemical reactions. *Journal of Research in Science Teaching*, 35(5), 569-581.

Kocakülah, S., Üstünlüoğlu, E., & Kocakülah, A. (2005). The effect of teaching in native and foreign language on students' conceptual understanding in science courses. *Asia-Pacific Forum On Science Learning And Teaching*, 6(2), Article 2.

Kruger, C. (1990). Some primary teachers' ideas about energy. *Physics Education*, (25), 86-91.

Kruger, C., & Summers, M. (1989). An investigation of some primary teachers' understanding of changes in materials. *School Science Review*, 71(255), 17-27.

Kruger, C., Palacio, D., & Summers, M. (1992). Surveys of English primary teachers' conceptions of force, energy and materials. *Science Teacher Education*, 76 (4), 339-351.

Küçük, M., Çepni, S., & Gökdere, M. (2005). Turkish primary school students' alternative conception about work, power and energy. *Journal Physics Teacher Education Online*, 3(2), 22-28.

Lee, O., Eichinger, D.C., Anderson, C.W., Berkhemer, G.D., & Blakeslee, T:D. (1993). Changing middle school students' conceptions of matter and molecules. *Journal of Research in Science Teaching*, 30(3), 249-270.

Lijnse, P. (1990). Energy between the life-world of pupils and the world of physics. *Science Education*, 74(5), 571-583.

Lin, C.Y., & Hu, R. (2003). Students' understanding of energy flow and matter cycling in the context of the food chain, photosynthesis and respiration. *International Journal of Science Education*, 25(12), 1529-1544.

Liu, X., & Tang, L. (2004). The progression of students' conceptions of energy: A cross-grade, cross-cultural study. *Canadian Journal of Science, Mathematics & Technology Education*, 4(1), 43-57.

Marek, E. (1986). They misunderstand but they'll pass. The Science Teacher***, 53***(9), 32-35.

Maskill, R., & Pedrosa, H. (1997). Pupils' questions, alternative frameworks and the design of science teaching. *International Journal of Science Education*. 29(7), 781-799.

Munson, B. H. (1994). Ecological misconceptions. Journal of Environmental Education***, 24***(4), 30-34.

Novick, S., & Nusbaum, J. (1981). Pupils' understanding of the partuculate nature of matter. *Science Education*, 65 (2), 187-196.

Özay, E., & Öztaş, H. (2003). Secondary students' interpretations of photosynthesis and plant nutrition. *Journal of Biology Education*, 37(2), 68-70.

Ross, K. (1993). There is no energy in food and fuels-but they do have fuel value. *School Science Review*, 75(221), 39-47.

Shepherd, D. I., & Renner, I. W. (1982). Students' understanding of the states of matter and density changes. *School Science and Mathematics*, 82, 650–665.

Stavy, R. (1991). Children's ideas about matter. *School Science and Mathernatics*, 91(6), 240-144.

Stevetrash Energy Pyramid, (2001). Retrieved December 20, 2006, from http://www.stevetrash.com/booking/lessons/pyramid.htm

Stylianidou, F., Ormerod, F., & Ogborn, J. (2002). Analysis of science textbook pictures about energy and pupils' readings of them. *International Journal of Science Education*, 24(3), 257-283.

Summers, M., Kruger, C., Mant, J., & Childs, A. (1998). Developing primary teachers' understanding of energy efficiency. *Educational Research*, 40(3), 311-328.

Swackhamer, G., & Hestenes, D. (2005.) An energy concept inventory. The ECI And BECI Are Online (Password Protected) At <Http://Modeling.Asu.Edu/Mns/Mns.Html>

Trumper, R., Raviolo, A., & Shnersch, A. M. (2000). A cross-cultural survey of conceptions of energy among elementary school teachers in training–emprical results from Israel and Argentina. *Teaching and Teacher Education*, 16, 697-714.

Trumper, T. (1997). Applying conceptual conflict strategies in the learning of the energy concept. *Research in Science and Tecnological Education*, 15(1), 5-18.

Vargo, L. A. (1997). Teaching energy to high school general biology students, Ph.D. Thesis, College of Natural Science, Michigan State University, U.S.A.

Watts, D. M. (1983). Some alternative views of energy. *Physics Education*, 21, 154-156.

ASSESSING TABULATING DATA SKILLS WITH ANALYTICAL RUBRICS

Burak Kağan Temiz
Nigde University, Faculty of Education

Abstract

This study is conducted to develop an assessment tool for assessing the constructing data table abilities of first grade students at high school. In the Tabulating Data Skills Test (TDST), the students are asked to organizing data into tables by using data provided in a text or a drawing. The answers of the students were assesses with the analytical rubrics developed. While developing the analytical rubrics, inter-rater and interior consistency analyses were done. The scope validity of the grading tool was checked by an expert. As a result of the study, reliable and valid analytical rubrics, which can be used in assessing the data tables at high school level, have been developed.

Introduction

Science process skills are thought skills where we use for establishing knowledge, thinking of problems and formulation of the results. These skills are the skills that scientists use during their researches. We can provide the students' learning and understanding their own world by making them acquire these important skills. These skills are base of thought and research which are included by science (Lind, 1998). Scientific processes are intellectual and physical skills that are utilized for collection of information, arrangement of the collected data under the light of different methods, explanation of extraordinary situations and solution of problems. Scientific processes are intellectual (and sometimes physical) activities that are performed when thinking, collecting data, and interpreting data or when utilizing the understandings achieved from data (Harlen, 1999). Science process skills have been called life long learning skills, as they can be used for daily living and for learning in school in any area (Carin, 1993).

Several authors identified and classified different sets of science process skills. In general; commonly used science contain several skills which may be stated as observing, classifying, measuring, ability to use numbers and space/time relationships, predicting, recording data, using the data for creating models, interpreting data, inferring, identifying and controlling variables, formulating hypotheses, and experimenting. Constructing a table of data skill has been approached by many researchers within the scope of data recording or communication skills. Constructing a table of data skill is within the scope of communicating skill and provides a basis for graphics drawing.

Data Recording is a skill regarding organization of the data, which have been collected for events and objects, under various regulatory forms used in scientific literature. While making experiments, students obtain both many qualitative and quantitative data. Measurements, which have been performed during a research, are called as data. Time, heat and volume measurements may be stated as examples of data. These data, which have been collected relating with events and objects, are recorded under various regulatory forms where everyone is easily able to understand. These regulatory forms facilitate utilization of data. (Hughes and Wade, 1993; Rezba, Sprague, Fiel, Funk, Okey & Jaus, 1995).

When an investigation is conducted, the measurements made are called data. Organizing data into tables helps to see patterns in the results. Rezba and colleagues (1995) stated that one of the skills needed to conduct an investigation is the organization of data in tables. When data are presented in well-organized tables, trends and patterns of change in data are often revealed. Before conducting a meaningful investigation, it is important to learn how to organize the data you have collected. By organizing data, a scientist can more easily interpret what has been observed. Making sense of observations is called data interpretation. Since most of data scientists collect is quantitative, data tables and charts are usually used to organize the information. Graphs are created from data tables. They allow the investigator to get a visual image of observations which simplifies interpretation and drawing conclusions. Valid conclusions depend on good organization and clear interpretation of data (Bailer, Raming & Ramsey, 1995). Scientists communicate orally, with written words, and thorough the use of diagrams, maps, graphs, mathematical equations, and other visual demonstrations. Graphs, charts, maps, symbols, diagrams, mathematical equations, and visual demonstrations, as well as the written or spoken word, are all methods of communication used frequently in science (Abruscato, 2000; Rezba et.al. 1995).

Rationale

Science in The National Curriculum (1991) asserts that, It is important for children to communicate their thoughts and ideas at all stages of their scientific activities and all levels of their development. "They should have the opportunity to express their findings and ideas to other children orally and through drawings, charts, models, actions and writing"(Hughes & Wade, 1993). Blackwell and Hohmann (1991) stated that children use language, pictures and mathematical symbols to present, and thus communicate, their observations and findings. Representation of observations and findings not only serves the practical function of sharing results with others but also helps children sharpen their observations and clarify their thinking. The data tables are the scientific communication tools extensively employed in specifying the products of the experiment. The artifices of developing a data table together with other scientific process skills are the substantial talents required for particularly laboratory work.

Data recording is an important skill, because; before conducting a meaningful investigation, it is important to learn how to organize the data you have collected. By organizing data, a scientist can more easily interpret what has been observed. Valid conclusion depends on good organization and clear interpretation of data. Additionally, organization of data under a table facilitates determination of relations among results as well as determination of patterns (Bailer et.al. 1995; Ostlund, 1992).

Methods

In this research; the name of the test, which has been developed in order to assess data table construction skills of students, has been defined as Tabulating Data Skills Test (TDST). During TDST development process; the procedural steps have been followed as shown in the Figure 1.

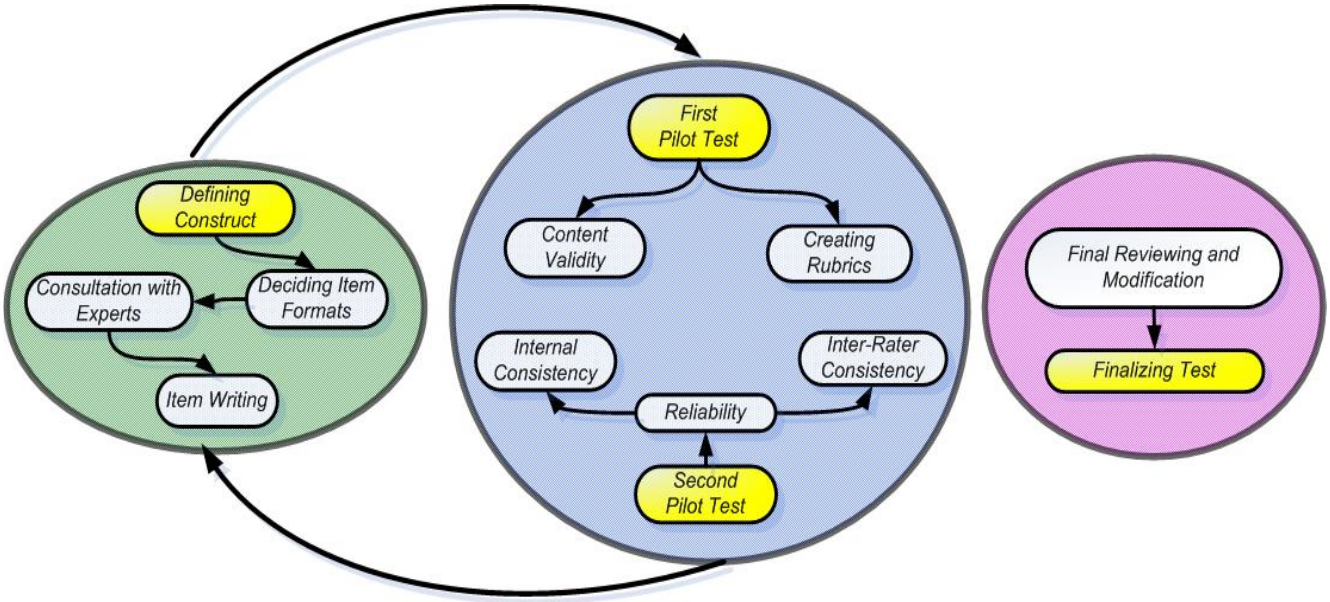


Figure 1. TDST Development Process.

The three main processes, which have followed during test development process, are: defining construct, pilot testing and finalization of the test (making the final amendments on the test).

Defining Construct

Performance assessment approach has been adopted during TDST development process. The following procedural steps outline the main factors to consider when making performance assessments (Gronlund, 1998).

Specifying the Performance Outcomes

During the first step, firstly, the performance outcomes, which will be measured during the test, have been determined. For this purpose, answers of the following questions have been searched;

What are the kinds of tables? The answer of this question depends on the number of variables which has been studied on and the number of recorded data. If the experiment, which has been performed, consists of one single step, two variables are in question. These are Manipulated variable and the corresponding responding variable. As the controlled variables have been kept constant throughout the experiment, there is no need to record them into the data table. If the experiment, which has been performed, consists of more than one step, more than two variables may exist. In each part of the experiment, the affect of a different manipulated variable on the responding variable may be tested. The data, which will be recorded at the end of the experiment or observation, may be quantitative data or qualitative (or descriptive) data. The data tables, which are necessary to be prepared for all these issues, may be different from each other. In other words, tables vary in accordance with the number and kinds of the variables to be recorded.

What are the rules which should be followed while drawing a data table? During the researches on the literature; some rules, which should be obeyed during preparation of a data table, have been determined. According to Rezba and colleagues (1995), Although there are no absolute rules for constructing tables of data, there are conventions, or commonly agreed rules upon patterns of organization, that facilitate communication between the writer and the reader. For example, when constructing a table of data, the manipulated variable recorded in the left column and the responding variable is recorded in the right column. Whenever units are used, they are included into the column heading as well. When recording data in a table, the levels of the manipulated variables are ordered. Although data are sometimes ordered from the largest to the smallest, the usual procedure is to order data from the smallest to the largest. This organization establishes a pattern of change in the manipulated variables. If there is a corresponding pattern of change in the responding variable, it will be easier to recognize if the levels of the manipulated variable were placed randomly in the table.

Selecting the Focus of the Assessment (Procedure, Product, or Both)

During TDST, students are expected to organize tabulation of the data groups which have been given in text or picture. Therefore, in TDST, it has been planned to make assessments on student performances by examining the tables drawn by them (in other words the product).

Selecting an Appropriate Degree of Realism

While table drawing skills are assessed during TDST, students have been requested to draw a real table on the space which has been left blank on the paper. The materials, which are necessary to answer TDST, consist of pen, rubber and ruler.

Selecting the Performance Situation

During TDST, students are expected to organize tabulation of the data groups which have been given in text or picture. Therefore; during TDST, students are requested to fulfill paper and pen performance tasks.

Selecting the Method of Observation, Recording and Scoring

During performance assessment, there are various methods for observation of student outcomes, recording the observation results and giving points to them. These methods are decided according to the characteristics of performance task and product, process or assessment or both. Some methods have been shown in the figure.

During TDST, analytical rubric has been utilized for assessment of student tables. Analytical rubric is the form which consists of performance criteria classified into various levels. By the analytical rubric, during assessment of data tables prepared by students, criteria may be formed by subcategories such as title of table, table structure, variable names, units, data recording and order. For this reason, the analytical rubric has been preferred for TDST.

Item Writing

During TDST; the task, which has been assigned to students, is to show the data in a data table as they have been given to them in text or picture. The questions, which exist in TDST, have been prepared for this purpose. The questions have been prepared in Turkish. The first version of TDST consisted of 20 items. Divisions of these 20 items according to tasks are given in the Table 1:

Table 1. The First Version of TDST

Tasks	Number of items
Tabulation of qualitative data with two variables	3
Tabulation of quantitative data with two variables	3
Tabulation of quantitative and qualitative data with two variables	3
Tabulation of quantitative data with more than two variables	3
Tabulation of qualitative data with more than two variables	3
Tabulation of quantitative and qualitative data with more than two variables	3
Table indicating preparations before an experiment or observation	2

Pilot Testing

During development of TDST, pilot tests have been realized. The first pilot testing has been realized in order to, give decision on the sufficiency of the answering times given for the test, prepare the test instruction, and determine the questions which students have difficulty in understanding. The first pilot test has been realized on 63 high school freshmen who receive education in 3 different schools located at the capital city of Turkey, Ankara.

The issues and observations, which have been determined during the first pilot test, are as follows: Great majority of students were not able to understand the task regarding preparation of the table indicating preparations before an experiment or observation (in other words, preparation of a blank table task), The number of questions were quite much and students were able to fulfill averagely 7.5 of 20 table drawing tasks which have been assigned to them. Answering ratio of the questions, which contain long texts, is low.

This fist pilot test has been performed in order to observe the conformity of the test with the level of student during implementation. Great revisions have been realized on TDST after the first pilot test. These are as follows: After the first pilot test, the answers of students have not been assessed and points have not been given to answers. Questions have been simplified to more simple structures.

Accordingly, only two types of tasks have been decided to be assigned during the test. These are as follows:

1st Type: contains tabulation (in other words tables with two variables) of data which have been collected resulting from the experiments consisting of one single step.

2nd Type: contains tabulation (in other words, tables with more than two variables) of data which have been collected resulting from the experiments consisting of more than one step. For some questions, it was decided to concentrate on figures showing the steps of experiment instead of long texts. The total number of tasks was decreased to 8. Accordingly, the content of the second version of TDST is given in the Table 2.

Table 2. The Second Version of TDST

Tasks	Number of items
Tabulating of the data collected resulting from the experiments which consist of one single step	4
Tabulating of the data collected resulting from the experiments which consist of more than one step	4

Opinions of physics teachers and academicians have been received for the second version of TDST. Thus, evidences have been collected for the content validity and face validity of the test. Experts have submitted their opinions for the second version of TDST by reading the test and taking notes on it, by filling the Test Assessment Form or orally during conversations.

The test has been reviewed again in accordance with the opinions of the experts and it has been made available for the second pilot test. Second pilot test has been realized in order to: decide on the sufficiency of the answering times given for the test and, utilize from the variety of answers while developing analytical rubric. (In fact, the development process regarding Analytical Rubrics to be utilized for assessment of TDST has been commenced earlier than the first pilot testing. Scoring categories and the criteria have become final by the second pilot testing. Preparation of analytical rubric has been postponed after the second pilot test in order to form a scale which is available for scoring for all the circumstances that may be faced with.)

The second pilot test has been realized on 33 high school freshmen who receive education in Ankara. Activities performed after the second pilot test: Analytical Rubrics have been developed for assessment of the 1st and 2nd type of tables. Reliability studies have performed for analytical rubrics.

Results

Reliabilities of the developed analytical rubrics have been searched by two methods. Inter-rater agreement analysis and internal consistency.

The reliability of the analytical rubrics, which have been developed to assess the TDST, has been tested through inter-rater agreement analysis. The student responses for the inter-rater agreement analysis have been evaluated by independent raters. For these assessments, the answer sheets of 20 students have randomly been selected amongst all students who have taken part in the second pilot test. These 3 raters have given points to totally 160 tables, which have been drawn up by 20 students, by using analytical rubrics. The points, which were given by the raters, have been compared with the Kappa Test results and their concordance rates have been scrutinized Kappa coefficient has been calculated by means of SPSS 11.5 packaged software. The results of this analysis are shown in Table 3.

Table 3. Inter-Rater Agreement Rates for Data Tabulation Related Analytical Criteria Scales

Recording Variable Number	Item No	Inter-Rater Agreement (Kappa)		
		Rater A- Rater B	Rater A- Rater C	Rater B- Rater C
Two variables	1	0.797	0.937	0.782
	2	0.799	0.937	0.784
	3	0.724	0.880	0.725
	4	0.808	0.855	0.780
More than two variables	5	0.801	0.967	0.784
	6	0.884	0.805	0.799
	7	0.857	0.860	0.823
	8	0.791	0.952	0.777

During TDST, students have been requested to fulfill a performance on paper. Each question requires fulfillment of a distinctive performance. For this reason, analyzes have been realized question by question while searching on the inter-rater agreement.

The weakest aspect of using an assessment tool, which contains open-ended items, is the subjectivity for assessment. Different assessments by different persons for the same product will affect the reliability of the scale negatively. The Kappa Statistics, which has been developed by Cohen (1960), is used for determination of the assessment agreement between/among two or more observers. If the Kappa coefficient is between 0.40 and 0.75, this means a reasonable agreement and if the coefficient is more than 0.75, this means an excellent agreement (Şencan, 2005). Within this context; the Kappa coefficients, which have been presented in the Table, indicate that inter-rater consistency reliability of the scoring items (tools) is high.

As for the internal consistency of the items being another part of the reliability studies has been achieved through computation of the Cronbach's Alfa Coefficient. To do so, the means points granted by the arbitrators according to the rubrics have been utilized. The Cronbach's Alfa Coefficient has been calculated to read 0.985 for the double-variable data, as for data possessing more than one variable 0.811.

Finally all these analyses indicate that the reliability of analytical rubrics, which are used for TDST assessment, is high in terms of both internal consistency and inter-rater agreement. As the statistics after the second pilot test indicate that TDST and Analytical Rubrics are in desired levels, and therefore, a new pilot test was not considered necessary and thus, the test has been finalized (last revisions have been made on the test).

Conclusions and Implications

When the literature is scrutinized, it appears that too many tests were developed for drawing a graphic and interpretation. Yet, any tool that shall assess the tabulating data skills has not been come across. During this study, an assessment instrument has been developed that might be used to assess the constructing data table abilities of the students. In order to effectuate the grading for the data table the students have prepared, two each analytical rubric have been developed and these rubrics have been tested from the standpoint of validity and reliability. First analytical rubric has been developed for the double-variable data tables, as regards to second rubric for assessment of the data tables with more than two variables. It is being envisaged that TDST might be employed during assessment of the data recording skills of the students during laboratory studies

In conclusion; valid and reliable an assessment tool, which may be utilized for assessment of high school students' skills for recording data into tables, have been developed by this study. The characteristics of the developed test are as follows:

There are 8 table preparation tasks in the test. Four of them are relating with tabulation of the data belonging to the experiments consisting of one single step and the other four tasks are relating with tabulation of the data belonging to the experiments consisting of more than two steps.

Implementation period of the test is one course hour.

The tables drawn up by the students are assessed with the developed analytical rubrics.

By analytical rubrics, points are given to student tables in accordance with categories such as Title of Table, Table Structure, Variable Names, Units, Data Recording and Order.

Analytical rubrics should be used for observation of students' table drawing skills, not for giving points to them. Thus, it may be possible to provide feedback to students for their mistakes.

TDST may be utilized for observation the skills' development in terms of recording the necessary data for laboratory studies in a table, for formative and diagnostic assessment (during instruction) or at the end of the instruction summative assessment.

References

Abruscato, J. (2000). Teaching Children Science, Needham Heights, M.A: Allyn and Bacon, 37-52.

Bailer, J., Ramig, J. ve Ramsey, J. (1995). Teaching Science Process Skills, Torrance: Good Apple Publication.

Blackwell, F. F. & C. Hohmann (1991). Science, High/Scope K-3 Curriculum Series. Ypsilanti, Michigan: The High/Scope Pr.

Carin, A. A. (1993). Teaching Science Through Discovery. Toronto: Macmillan Publishing Company 3-17.

Grounlund, Norman E. (1998). Assessment of Student Achievement. Allyn & Bacon.

Harlen. W. (1999). Purposes and procedures for assessing science process skills. Assessment in Education: Principles. Policy and Practice. 6(1), 129-145.

Hughes, C. & Wade, W. (1993). Inspirations for Investigations in Science. Warwickshire: Scholastic Publication.

Lind. K. (1998). Science Process Skills: Preparing for the future. Monroe. Available: http://www.monroe2boces.org/shared/instruct/sciencek6/process.htm [Retrieved 1998. June 15].

Ostlund, K. L. (1992). Science Process Skills: Assessing Hands on Student Performance. California: Addison Wesley.

Rezba, Richard J., Ronald L. Fiel ve H. James Funk (1994) Learning and Assessing Science Process Skills. Kendall Hunt Pub.

Şencan, H. (2005). Sosyal ve Davranışsal Ölçümlerde Güvenirlik ve Geçerlilik. Ankara: Seçkin Press.

RETAINING WEAKER STUDENTS: A PILOT PROJECT IN CHEMISTRY AT THE UNIVERSITY OF LIMERICK

Sarah Hayes
University of Limerick

Peter E. Childs
University of Limerick

Abstract

In Ireland many universities, including the University of Limerick have a problem with weak students in undergraduate Science programmes leading to high failure rates. Entry to these courses is possible with low grades in Mathematics and no relevant Science subject from second level. Those who have not done a specific Science subject at second level are often left behind, as they lack the necessary Science background. The Chemistry module that was examined typically has a failure rate of 30-40% at the first sitting. It is taken by students from five different degree courses (A-E). Students who fail this module are usually those in course 'A' and course 'B', and they are significantly weaker on entry than those in courses C, D and E. Therefore we designed a 9 week intervention programme for groups A and B based on pre- and post-concept and attitude diagnostic tests. Results based on this intervention show, on average, a positive trend in both conceptual understanding and confidence levels. Over 76% of students who took both pre- and post-diagnostic concept tests improved in their performance. The performance of these two groups has also been evaluated in the final examination results, in two modules compared to their peers who did not take the programme.

Introduction

Since the eighties third level education in Ireland has undergone considerable expansion. This increase in numbers has been largely due to the government's policy on higher education. In particular, the last ten years has seen the largest increase in numbers with 50-55% of 17-18 year olds entering higher education in 2007, with the goal being to reach 72% by 2020. (Higher Education Authority, 2008)

This rise in numbers means that there is a very diverse group of students in the third level classroom. (Childs, 2009) (Darmody and Fleming, 2009) As Zeegers and Martin (2001, p.35) note *'gone are the days when university classes contained only highly selected students, with present day classes now containing students with a more diverse range of academic skills, past teaching and learning experiences, prior knowledge, approaches to learning and expectations of the tertiary experience*" and this is an issue which is even more pronounced now. As a consequence of increased participation rates both in Ireland and worldwide, there has been an increase in student failure rates and student attrition. The Higher Education Authority study on completion in Ireland (Morgan, Kellaghan and Flanagan, 2000) noted that Irish students do tend to have better 3rd level completion rates than their European counterparts. Science and Mathematics across Europe have higher non-completion rates than other subjects, such as Science and Mathematics. Science courses in Ireland, at third level have considerably lower rates of non-completion (22.2%) in comparison to other courses such as Law (7.1%). (Moore, 2004) (Flanagan and Morgan, 2004)

This has led Irish Universities to look closely at the subject of retention identifying *'retention, completion and student withdrawal as important issues to be addressed'* (Moore, 2004). These issues have been receiving increasing attention over the last three years. In particular it was noted that there was a need for learning support programmes for students who were 'weak' in critical areas such as Science and Mathematics. (Flanagan and Morgan, 2004)

Rationale

There is increasing diversity of undergraduate ability, which leads to a high student attrition rate. Entry into Irish third level institutions is based on the number of points accrued in the Leaving Certificate examination. The Leaving Certificate is a state examination, which 2nd level students take at the end of their second level education, usually aged 17-18.

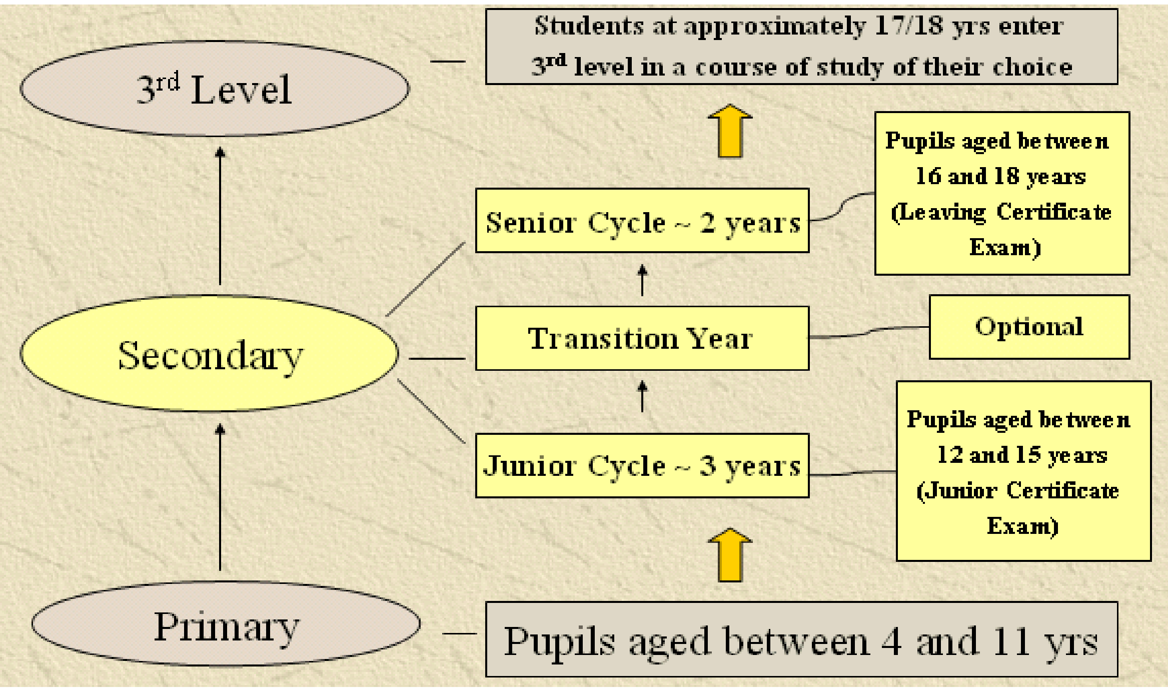


Figure 1. The Irish School System (Childs and Sheehan, 2009).

Students can take as many subjects as they wish, but points for entrance into third level are taken from the best six subjects. An A1 grade at Leaving Certificate level earns 100 points and a D3 grade earns 45 (on a higher level paper). A maximum of 600 points can be earned through this system. Course entrance points for each year are determined based upon the number of course places and the student demand for the courses. Therefore the number of points that one needs to be accepted into a course does not necessarily reflect the difficulty of a course, but its popularity. The Leaving Certificate examination results, third level applications and offers of course places from third level institutions are all processed through the Central Applications Office (CAO). This entrance system leads to many courses, in particular, Science courses, having low entrance points and consequently, high failure and attrition rates, as the students are choosing more highly valued professional courses such as Law, Medicine and Pharmacy.

Students entering third level Science courses are not always required to have completed a relevant Science subject at Leaving Certificate level and these students are very often left behind. (Childs and Sheehan, 2009) It has found this to be a particularly crucial issue in Chemistry classes at third level. *'Chemistry is taken by 10-15% of students in the senior cycle of school (Leaving Certificate) in Ireland each year, and therefore tertiary institutions cannot impose a prerequisite of chemistry for entry into chemistry based degrees because of the limited pool of potential applicants'* (Seery, 2009) Those students who have not done Chemistry at school, do not have an adequate grounding in the basics of Chemistry.

The 2nd year Chemistry module (CH4253) that we examined has an average failure rate of 30 – 40% at the first sitting. The module is taken by students from five different degree courses (A – E) with entrance points for these courses varying from 325 - 485. The two courses 'A' and 'B' accounted for 95% of the failures in the year 2007. One reason for this may be that the cut-off points for these two courses are lower than those of the other courses taking the module. Also, there are a larger percentage of mature students in these two groups, with fewer having completed Chemistry at Leaving Certificate. A large number of mature students in a class group can have implications for the entire class as they, typically have a more varied background than the rest of the 18-22 year old class cohort. (Kelly, 2005)

This has led to the following research questions to guide this study:

- Can diagnostic concept tests, identifying students' prior chemical knowledge and misconceptions, be used to design successfully an intervention programme?
- How do students' attitudes and confidence in the area of Chemistry affect their relationship to performance?
- Does this targeted intervention have an effect on the students' performance in the pre- and post-concept tests?
- Does attendance at the intervention programme made a difference in students' overall performance on the course?

Methods

In order to promote chemical understanding among the students in courses 'A' and 'B' a pilot intervention programme was developed. The programme was designed for the two groups of students in the first academic semester of 2008-09. A nine week programme of specialized tutorials covering basic chemistry concepts that had proven to be problematic was offered to the students. This was an optional programme and no extra course credits were offered to those that did. The most obvious limitations of this were that we were only able to measure the performance of those students who attended most of the sessions, and for those who took the pre- and post- test.

We know that there are *'flaws in the standard approach'*. (Herron, 1999, p.3) The programme was designed based on the results of a diagnostic pre – test that students were given in the first session. This was done in order to meet the students' needs, rather than attempting to address student understanding with our own pre – conceptions of what chemical concepts the students found difficult. We aimed to teach smarter, rather than to teach harder. See Figure 2 below. (Perkins, 2007)

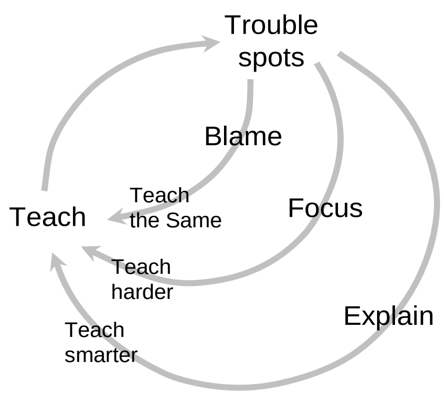
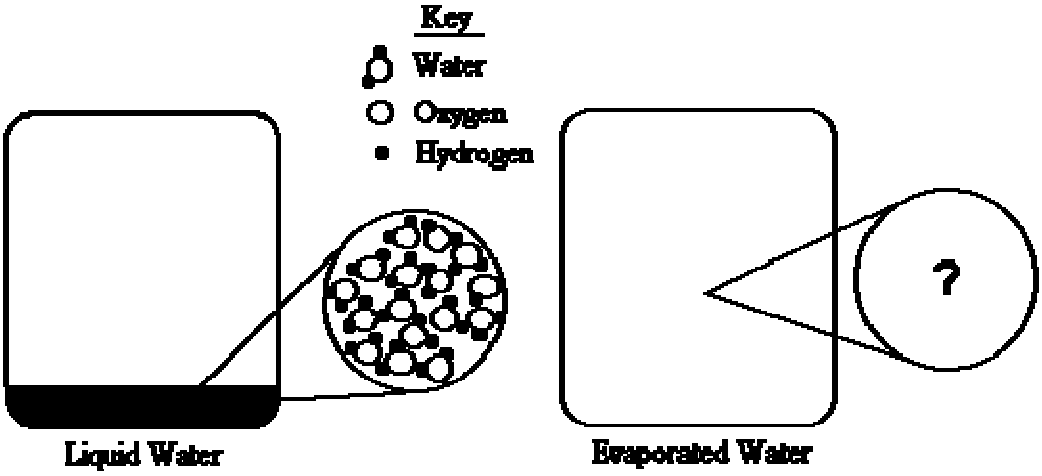


Figure 2. Three responses to trouble spots in learning (Perkins, 2007).

The pre - and post - diagnostic test of chemical concepts and misconceptions was designed and administered in the first and last tutorial session. Questions were taken from various chemical concept inventories and the test was designed to suit our needs. (Journal of Chemical Education, 2008) (Mulford, 1996) (National Institute for Science Education, 2008).A sample chemical concept question is shown in Figure 3.

The circle on the left shows a magnified view of a very small portion of liquid water in a closed container.



What would the magnified view show after the water evaporates?

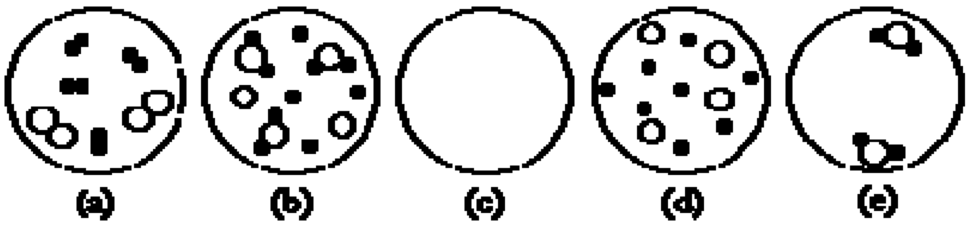


Figure 3. Concept test sample question on both pre- and post-tests.

The test also included an instrument measuring student attitudes and confidence towards chemistry. These attitude tests had a six point Likert scale, with a N/A option. Students from the two courses were taken in separate small class groups, in a tutorial style format. This was done in order to meet the differing needs of the two classes. The unique feature in this study was to use the pre-chemical concept diagnostic test results to design the intervention programme for the students, in order to address their misconceptions and lack of understanding.(Anders and Berg, 2005) As we know *'many chemical processes are familiar in everyday experience, the chemical concepts underlying these processes are themselves unfamiliar'* (Gabel, 1999, p.3) It is important to note that the diagnostic pre- and post- tests were designed to diagnose students' chemical misconceptions, rather than just their chemical knowledge.

Results

Overall we have experienced positive results in this intervention study. Students were given a pre- and a post- test of confidence/attitude. Student's confidence to understand key concepts of chemistry improved in both groups; Group 'A' had an improvement in confidence levels and went from 50%, of respondents choosing average and high, to 73% choosing those two options. Group 'B' had their confidence levels go from 50%, choosing average, to 56% choosing average and high. (A shift of one point on the Likert scale) However, overall improvement in confidence levels in both groups was not significant. I.e. Group 'A' and Group 'B' ($p > 0.05$)

Pre – and Post – Diagnostic tests of chemical concepts

There was a significant difference between pre – and post- concept test results for both groups, in a positive direction. See Figure 4.

On average participants in Group 'A' experienced significantly higher results in the post-concept test after the intervention programme (M = 39.08, SE = 4.62) than in the pre-concept test (M = 29.73, SE = 4.17, $t(11) = -2.94$, $p = 0.014$, $r = 0.66$).

Participants in Group 'B' also experienced significantly higher results in the post-concept test, after the intervention programme (M = 41.28, SE = 4.49) than in the pre-concept test (M = 22.05, SE = 2.86, $t(12) = -3.80$, $p = 0.003$, $r = 0.72$)

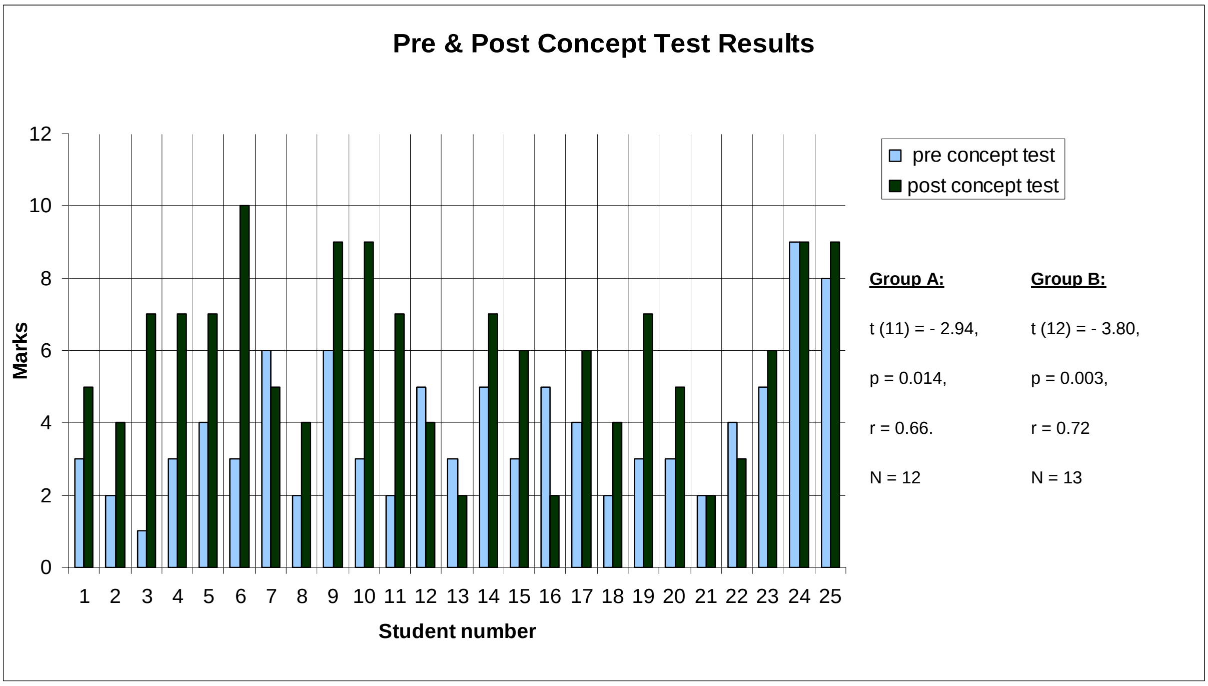


Figure 4. Comparison of Pre- and Post-test results for groups 'A' and 'B'.

In course 'A', 30% of those who completed both the pre- and post – diagnostic test had Leaving Certificate Chemistry. In course 'B' the numbers were lower, with 15.38% of students who had completed both the pre- and post- test stating that they had Leaving Certificate Chemistry. This was considered to be significant for both groups $p < 0.05$ in the results of the pre and post chemical concept test.

Intervention Programme impact on further Chemistry Modules

However, it was expected that students would improve in the pre- and post- concept test. Students in groups 'A' and 'B' take two chemistry modules in their second year. In order to examine whether the intervention programme had a more lasting effect, it was decided to examine students' results in the Chemistry module that was running alongside the intervention programme and in students' Chemistry module in the second semester, taking by the same semester cohorts. The impact of the intervention programme on the results of these modules can be seen in Figure 5.

On average, in module CH4253, students in both groups that took part in the intervention programme did better than their peers who did not.

In group 'A' participants in the intervention programme experienced slightly higher grades (M= 47.63, SE = 3.38) than those who did not take part in the intervention programme (M = 40.21, SE = 3.22). This difference was not significant t (29) = 1.53, p = 0.138; however it did represent a small to medium size effect r = 0.27.

Participants of the intervention programme in group 'B' experienced much better grades (M = 34.09, SE = 2.51) than those in the group who did not take part in the intervention programme (M = 21.84, SE = 3.47). This difference was considered highly significant t(25) = 2.82, p = 0.009; there was also a medium -large size effect r = 0.49.

There was thus a much greater 'value-added' effect for the weaker group of students, as one might expect. It was decided to also evaluate the performance of the two groups in a subsequent module (CH4554), taken in the second semester of second year. Again the performance of the weaker students improved much more so than that of the more able students, as shown in Figure 6.

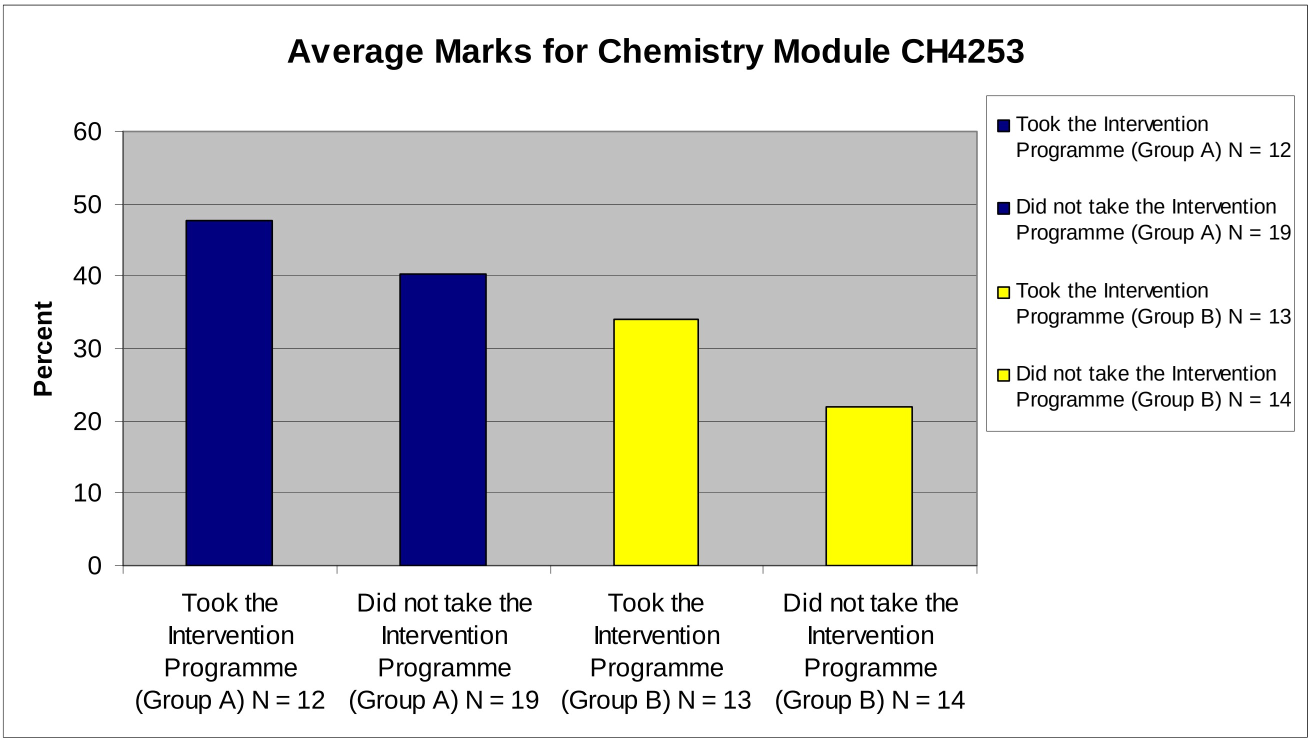


Figure 5. Comparison of results in module CH4253.

In the chemistry module CH4554 students who took part in the intervention programme were also seen to have better grades than those who did not take part in the programme.

Students in group 'A' who participated in the intervention programme had higher grades (M = 47.64, SE = 2.29) than those who did not participant in the intervention programme (M = 38.35, SE = 5.79). This difference was not significant t (23) = 1.50, p = 0.155; however it did represent a medium size effect (r = 0.3).

Those in group 'B' who took part in the intervention programme had better grades (M= 33.29, SE = 2.17) than those who did not take part in the intervention programme (M= 26.45, SE = 1.40). This difference was considered to be significant, t (22) = 2.64, p = 0.016. A medium -large size effect was also noticed (r = 0.49).

Again the weaker group (Group 'B') experienced a greater benefit from the programme, with students who took the intervention programme averaging a D1 grade (34.25%), in comparison with their peers who didn't take the programme averaging an F grade (21.84%)

Overall, there was a lower class failure rate in Chemistry Module CH4253 in the year 2008 (33.3%) (after the intervention programme), than in 2007 (41.6%). 82% of the fail grades in this module in 2008 were achieved by students in Group 'A' and Group 'B', compared to 95% in 2007. Most of the failures were in the non-attending group in both modules.

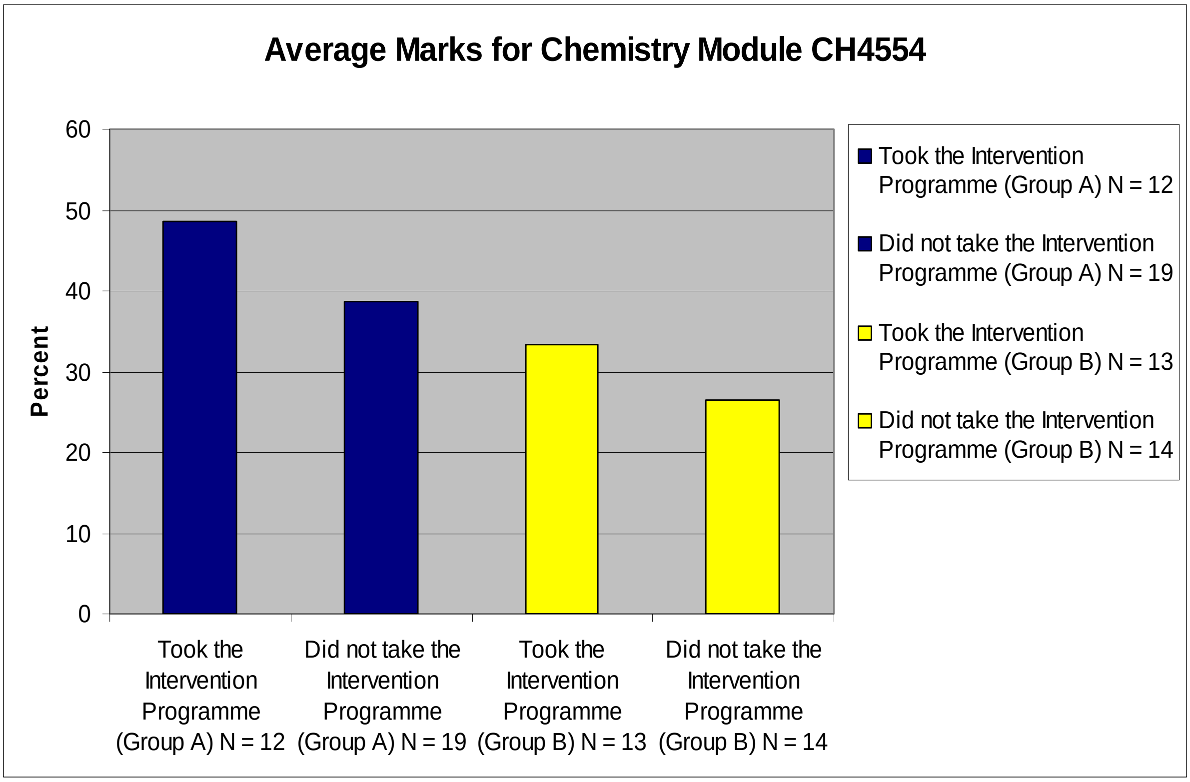


Figure 6. Comparison of results in module CH4554.

Conclusions and Implications

The results, in all areas, of the intervention programme are positive. Confidences in certain areas of Chemistry and examination results have all improved, as well as students' understanding of chemical concepts. However, this was an optional programme, and while the results are encouraging, poor attendance, both in the main modules and in the intervention programme does affect the results. Many students who attended the intervention programme could not be assessed as they did not attend both the pre- and the post- diagnostic test sessions.

The programme was offered to students in their second year of study. We intend to run a similar programme for the same groups of students in their first year of study, in order to equip them from the beginning of their course, with the basic chemical understanding that they need. The course will also be run over an extended time frame, for two semesters rather than one.

Further study needs to take place, in order to determine whether variables such as gender, CAO points obtained, course attendance, previous study of Chemistry or Mathematics, and overall college semester results have any effect on the results of the programme.

The implications of this study are that weaker students, doing science courses, who have lower entrance points and a poor background in Chemistry, are at a significant disadvantage when compared with their peers, who have a better background in Chemistry. Rather than lowering the standard of our third level courses, to allow for weaker students, it is important for us to teach smarter (Perkins, 2007) in order to help students adjust to third level and to make up for deficiencies in their background.

References

Anders, C., and Berg, R. (2005), 'Factors related to observed attitude change toward learning chemistry among university students', *Chemistry Education Research and Practice*, 6 (1), 1-18

Childs, P. E., and Sheehan, M., (2009), 'What's difficult about chemistry? An Irish perspective', *Chemistry Education Research and Practice*, 10, 204-218.

Darmody, M., and Fleming, B., (2009) 'The balancing act' - Irish part-time undergraduate students in higher education, Irish Educational Studies, 28(1), 67-83.

Flanagan, R. and Morgan, M. (2004) 'Evaluation of Initiatives Targeting Retention in Universities: A Preliminary Report of Projects Funded by the Higher Education Authority', 1st ed., Dublin: Educational Research Centre.

Gabel, D., (1999) 'Improving Teaching and Learning through Chemistry Education Research: A Look to the Future', Journal of Chemical Education, 76, 548-554.

Herron, J. D., (1999) 'Chemical Education Research: Improving Chemistry Learning', Journal of Chemical Education, 76, 1353-1361.

Higher Education Authority, (2008) *National Plan for Equity of Access to Higher Education 2008-2013*, Dublin: National Office for Equity of Access to Higher Education, Higher Education Authority

Journal of Chemical Education, (2008) '*Journal of Chemical Education QBank*', [online] available: http://jchemed.chem.wisc.edu/JCEDLib/QBank/collection/CQandChP/index.html [accessed on 4 Sept 2008]

Kelly, M., (2005) 'The effects of increasing numbers of mature students on the pedagogical practices of lecturers in the institutes of technology', *Irish Educational Studies*, 24(2-3), 207-221

Moore, S. (2004) 'Inter-Universities Retention Network-A Submission to the OECD Review Team on the Irish Higher Education System in Irish Universities', [online] available: http://www.iua.ie/publications/documents/publications/2004/Report_2.pdf [accessed on 27 May 2009]

Mulford, D. R.,(1996) '*An Inventory for Measuring College Students' Level Of Misconceptions in First Semester Chemistry*', unpublished thesis (M.S.), M.S., Purdue University.

National Institute for Science Education (2008), Field Tested Learning-Assessment Guide, *Tools,* [online], available: http://www.flaguide.org/tools/tools_technique.php [accessed on 2 Sept. 2008]

Perkins, D., 2007, 'Theories of difficulties', ch. 3 in *Student Learning and University Teaching,* BJEP Monograph Series II, (31-48), Leicester: The British Psychological Society

Seery, M., K., (2009) 'The role of knowledge and student aptitude in undergraduate performance in chemistry: a correlation prediction study', *Chemistry Education Research and Practice*, 10, 227-232

Zeegers, P., and Martin, L., (2001) 'A Learning-to-learn Program in a First-year Chemistry Class', *Higher Education Research and Development*, 20 (1), 35-52

STUDENTS' OPINIONS ABOUT WRITING JOURNAL IN SCIENCE-TECHNOLOGY AND MATHEMATICS LESSONS

Dilek Erduran Avcı
Mehmet Akif Ersoy University

Hasan Uslu
Ministry of Education

Abstract

The purpose of this study was to investigate the student opinions about writing journals in science-technology and mathematics lessons. This research was applied to 15 sixth grade and 15 seventh grade students in the 2007-2008 fall term. Research lasted nine weeks. During the research period, students wrote journals two times a week about subjects taught in those weeks. Finally, individual interviews were done with 27 voluntary students. Eight questions were asked to students about writing journals. Students define journals in interviews as an instrument that provides connection to the teachers, helps understanding lessons, helps reviewing and raising the exam marks and also contributing to performance and knowledge completion in lessons, and entertaining. According to students, journals contributed to create opportunity repeating lessons, to increase participation to lessons. They said journals also helped to exams, made them use punctuation marks carefully, suit spelling rules, write tidy and neat and improved their thinking, communication and confidence feeling in their relation to teacher. A lot of students expressed that journals helped exploring themselves. In this scope, journals can be used to help students express feelings and ideas, repeat the knowledge gained in the lessons and build up better student-teacher communication in lessons.

Introduction

Assessment is an integral part of educational activity (Kaptan, 1998). Alternative assessment methods which have a short history have been discussed by instruction and educational communities in many countries as well as Turkey. These have taken their place in science and technology institutional programmes (MEB, 2006) since 2004-2005 time periods in our country. This program is based on alternative measurement and evaluation approach based on the constructivist learning theory.

Alternative assessment, unlike the traditional approach, is identified as a student-centered, meaningful, interesting and appropriate student assessment strategy (Llewellyn, 2002). Alternative measurement and evaluation techniques are used for not only the evaluation of products but also the learning process. Therefore it makes the students feel the responsibility of learning and allows them to be proud of what they learn (MEB, 2006). Alternative assessment emphasizes testing and evaluation of different approaches and materials that are used to develop the problem-solving and scientific procedure. Another advantage of this approach is, being able to handle the developments of learning in cognitive, sensitive and psychomotor dimensions simultaneously (Çepni, 2005).

Student Journals

Science and technology journals are usually considered as the records of the activities performed in the course, kept by the students. Also known as science and technology notebooks, these journals encourage students to write their daily natural course experiences. Students can express the problems they try to solve, methods they use, observations they make and conclusions they reach via these journals. Journals have different types, such as thinking journals, affirmational dialogue journals, think-aloud journals and team journals (Ruiz-Primo, Li, Ayala & Shavelson, 2004). Journals also create an opportunity for students to express personal ideas and observations after their own experiences. Journal pages can be considered a story in which the observed events change by time (Shepardson, 1997, cited in Korkmaz, 2004).

There has been a consensus that student journals are progressive measurement instruments for the instructors. These journals allow instructors evaluate students' conceptual and procedural understanding and get required feedback to improve performance. Student journals also give information not only about learning of the students but also about the quality of the education. Through the multi-assessment approach, they can also be considered as quick evaluation instruments. Journal records are constructed by evaluation applications that match the purpose, content and instruction activities in the education process (Ruiz-Primo, Li,Ayala & Shavelson, 1999, March).

Student journals, which can be used as an information resource for student performance and instructor feedback (Ruiz-Primo, Li, Ayala & Shavelson, 2004), have several uses:

- To provide proof of student learning and progress to the instructor.
- To grant an opportunity to return personal feedback to students and plan personal education with the instructor.
- To grant an opportunity to both the instructor and students to create a dialog about the principals of the course and student progress.
- To make the instructor and students build up better personal communication.
- To improve students' writing and illustrating abilities.

Rationale

There are several studies in fields such as geography (Cook, 2000), medicine (Goldenhar and Kues, 2006), nursing (Barton and Brown, 1992) and science education (Ruiz-Primo, Li, Ayala & Shavelson, 1999, March ; Ruiz-Primo, Li & Shavelson, 2002; Shepardson & Britsch, 2001; Ruiz-Primo, Li, Ayala & Shavelson, 2004; Nesbit, Hargrove, Harrelson & Maxey, 2004; Hargrove & Nesbit, 2003) which were made by student journals. However, it is interesting that less work has been performed about this issue in science and math education especially in Turkey (Tektaş-Hasanoğlu, 2002; Erduran-Avcı, 2008). In addition, no study, using student journals in multiple courses during the same study period, has been encountered.

In contemporary primary school programs of Turkey, in addition to evaluate student achievement by grades, monitoring the development of individual specialties during the learning-teaching process and evaluation of students in the process of active participation gained importance. Student journals are very effective tools within the scope of individual assessment. In this context, this work, which is about the students' opinions on use of student journals in science and technology and mathematics courses, is likely to contribute to the literature.

Purpose

The purpose of this study was to investigate the student opinions about writing journals in science-technology and math courses.

Methods

This is a qualitative research. Research was performed in a primary school located in Tokat, Turkey, during the fall term of the 2007-2008 educational year, with the contribution of fifteen seventh grade and fifteen sixth grade students. The study lasted 9 weeks. During the research period, students wrote two journals in science-technology lessons and two journals in math lessons in a week. Thus, students wrote totally 36 journals in lessons about a course topic-related subject chosen by the teachers. The rest of the journal subjects were prepared similarly about the topics in lessons. Science-technology and math teachers read the journals and gave the students feedback about their journals. Finally, individual interviews were performed with 15 sixth grade and 12 seventh grade volunteer students. The interview is performed in a semi-formal form. In order to help the researcher perform interviewing and listening activities better, the meeting was recorded to camera. The analyses of data obtained from the interviews were done through the use of descriptive analyses.

Results

This section covers the answers of the students to the semi-structured interview questions which are prepared by the researchers. Individual interviews with students were carried out within the scope of 8 main questions. Multiple student opinions among the answers of a single question are included in the analysis.

Question 1. What are the science-technology and math journals? How can you define it?

Students defined these journals as tools that help them in repeating what they have learned, help them understand the lesson, keep records of the things that they did in the lesson and make the communicate with the teacher. Percentage and frequency values for answers to the first question are presented in Table-1.

Table 1. Student opinions about first question and percentage-frequency values

What are the science-technology and math journals? How can you define it?					
Codes	6th grade		7 th grade		Examples of student expressions
	f	%	f	%	
They are instruments for repeating the lesson.	8	53,3	2	16,6	-*"It is a production of things we did in math and science technology courses."*
They are instruments that help us understand the lesson.	5	33,3	5	41,6	-*"I write what we did in the course to my journal."* -*"It helped me contact my teacher."*
They are instruments that keep records of the lesson.	6	40	5	41,6	-*"What we have learnt in the lesson today, what we have done. It is used for repeating the lessons"*
They are instruments that make us communicate with the teacher.	-	-	1	8,3	-*"An activity to make us our lessons better."*

Question 2. Have you ever written journal?

According to the answers of students, it seems that they sometimes write journal, but not regularly. Many of their journals are about the daily life and include topics about memories. Percentage and frequency values for answers to the second question are presented in Table-2.

Table 2. Student opinions about second question and percentage-frequency values

Have you ever written journal?					
	6th grade		7th grade		
Codes	f	%	f	%	Examples of student expressions
No	3	20	3	25	-*"No, I have not"* -*"Yes. Once. I write now."* -*"I tried to. But I can not."* -*"I write about my daily life"* -*"I write about my memories."*
Yes	12	80	9	75	

Question 3. From what point of view these journals contributed to you?

According to the answers of the student's, journals contribute to students in repeating lessons, having a better understanding of lessons, not forgetting the subjects, achieving success in courses, improving thinking and reading, using mental capability, observing the life from different points of view, helping communication, improving writing, complying to the punctuation and spelling rules, and taking care of the order of the writing. Percentage and frequency values for answers to the third question are presented in Table-3.

Table 3. Student opinions about third question and percentage-frequency values

From what point of view these journals contributed to you?						
		6th grade		7th grade		
	Codes	f	%	f	%	Examples of student expressions
Lesson understanding	Made me have a better understanding of the lesson	4	26,6	6	50	-*"…..made me repeat my lessons, made me read, write and organize the paper in a regular way."* -*"...I cannot write slain to the journals. It contributed me in writing and changed my view of life. I looked some things from different viewpoints and learned better".* -*"..it helped me communicate better… I cared for punctuation and used my thoughts…"* -*"..it helped me improve my writing, achieve success, use punctuation marks, and care about order and organization."* -*"Made me understand the topics better and improved my thinking"* -*"…made me use my mind."*
	Made me not to forget the lesson	4	26,6	2	16,6	
	Made me repeat my lessons	6	40	2	16,6	
	Helped me in achieving success	6	40	1	8,3	
	Improved my reading	4	26,6	2	16,6	
Thinking	Improved my thinking	4	26,6	5	41,6	
	Made me observe the life from different viewpoints	-	-	1	8,3	
	Made me use my mind	1	6,6	2	16,6	
	Helped me in communication	-	-	1	8,3	
Writing	Improved my writing	13	86,6	8	66,6	
	Made me use punctuation marks carefully	4	26,6	4	33,3	
	Made me careful of organization and order	1	6,6	5	41,6	

Question 4. Is there any difference between daily life journals and these journals?

According to the answers of students; the difference between the daily journals and school journals is that school journal topics are associated with courses while the other journals are about daily life. Percentage and frequency values for answers to the fourth question are presented in Table-4.

Table 4. Student opinions about fourth question and percentage-frequency values

Is there any difference between daily life journals and these journals?							
			6th grade		7thgrade		
		Codes	f	%	f	%	Examples of student expressions
There are differences.	Subject scope	School journals are about lessons other journals are about daily life.	10	60	10	75	-*"Yes there are. School journals are about lessons. The others at home are related to daily life."* -*"Yes there are. The journals we write in lessons help us repeat what we have learnt."* -*"There are for sure. The journals I write outside are about my affairs with my friends. School journals are about the thing we learn in lessons."* -*"...I care about punctuation and order in the journals I write at school."* -*"...there are differences related to the thing I lived that day."* -*"In the journals about our daily life, I write the days which are special and meaningful for me. In school journals, I write about lessons and my experiences in the school."* -*".....no there are not. I started all with hello."*
	Writing rules scope	I care about punctuation marks and order in student journals.	1	6,6	1	8,3	
	Objective scope	Student journals are written for repeating what we have learnt.	2	13,3	1	8,3	
There are no differences.			1	6,6	-	-	

Question 5: What did you consider while writing a journal? Has there been any change in thoughts that you needed to watch out for?

Students took care of writing blogs and content issues, the spelling rules, punctuation marks, order and organization and writing the information they kept in mind about the lesson. Students stated that by the use of the journals, their writing, reading and thinking abilities improved and in addition, their knowledge of punctuation became better. Percentage and frequency values for answers to the fifth question are presented in Table-5.

Question 6: What can be the purpose of the journals that you wrote according to you?

Students have expressed the purpose of journals as: repeating courses, expressing ideas and feelings, helping them discover themselves, change their points of view, improve their handwriting, making them not to forget the topics covered in the lesson, making student achievement higher, increasing their motivation to the course, helping in exams increasing success and improving reading ability. Percentage and frequency values for answers to the sixth question are presented in Table-6.

Table 5. Student opinions about fifth question and percentage-frequency values

What did you consider while writing a journal? Has there been any change in thoughts that you needed to watch out for?

		Codes	6th grade f	6th grade %	7th grade f	7th grade %	Examples of student expressions
While writing the journal,	Writing	I cared for writing beautifully.	8	53,3	5	41,6	-*"I cared for subject and style."* -*"Generally I tried to write about course topic and content. I focused on the subject handled that day and expressed my thoughts."* -*"...I cared for punctuation."* -*"I cared for my writing, reading and keep the things in my mind."* -*"order and organization…meanings related to the topic of the lesson."* -*"…I care for the organization of the paper and meaning of my writing."* -*"I cared for spelling rules, use of punctuation, relations between topics and figures."*
		I cared for spelling and punctuation rules.	10	66,6	8	66,6	
		I tried to write in an organized and arranged way.	2	13,3	4	33,3	
		I tried to write meaningfully.	3	20	6	50	
	Subject / content	I cared for writing about subjects of the lesson.	3	20	4	33,3	
	Form	I cared about illustrations.	-	-	1	8,3	
Yes, there have been changes in me…		My writing improved.	9	60	7	58,3	-*"…improved my writing and I think more than I used to."* - *"Improved my writing and reading."* .-*"My handwriting improved, I learned the punctuation marks.* -*"… it improved my writing and thinking. It also improved our lessons and progress."*
		My reading improved.	2	13,3	1	8,3	
		I learned punctuation marks.	4	26,6	3	25	
		My thinking improved.	1	6,6	2	16,6	
		My lessons improved.	5	33,3	1	8,3	
		My in-class contribution improved.	1	6,6	-	-	

Table 6. Student opinions about sixth question and percentage-frequency values

What can be the purpose of the journals that you wrote according to you?

Codes	6th grade f	6th grade %	7th grade f	7th grade %	Examples of student expressions
To repeat lessons.	6	40	5	41,6	-*"...it can be making us repeat lessons. If we cannot understand in lesson, we understand better in journals."* -*"To make us express our feelings and ideas, to change our point of view to life and to make us discover ourselves."* -*"It is making us not to forget the things we did, we cannot keep the things we do in mind, if we keep notes, we can check every day."* -*"Because it makes us study and willing."* -*"To make me successful in the lessons. It can be for enhancing my grades in the exams."* - *"…it improved my reading. That can be."*
To express our ideas and opinions.	1	6,6	1	8,3	
To discover ourselves	-	-	1	8,3	
To change our view of life.	-	-	1	8,3	
To improve our writing.	-	-	1	8,3	
Not to forget the things we did.	5	33,3	4	33,3	
To gain a better understanding of lessons.	5	33,3	5	41,6	
To motivate us.	1	6,6	-	-	
To make us successful.	3	33,3	-	-	
To improve reading.	1	6,6	-	-	

Question 7. Can journals be a good communication tool between teachers and students? Were you being able to express things in journals that you wanted to share with teachers, but in everyday events could not?

A great majority of students agree that journals are good communication instruments between students and teacher. In addition, it has contributed to the development of self confidence via expressing their thoughts by writing that they could not say in the classroom. Percentage and frequency values for answers to the seventh question are presented in Table-7.

Table 7. Student opinions about seventh question and percentage-frequency values

Can journals be good communication tools between teachers and students? Could you express things which you wish to share with your teacher but cannot do in daily life, by using journals?

	Codes	6th grade f	6th grade %	7th grade f	7th grade %	Examples of student expressions
Yes, it can be a good communication tool…	It helps me in getting in contact to my teacher.	3	20	2	16,6	-*"Yes can be. We write a subject to our journals if we could not understand and express it."* -*"…it helped me to get in contact with my teacher regularly".* -*"I wrote my complaints."* -*"We can define ourselves to our teacher better. We can make ourselves known to our teacher better and gain a better knowledge of our teacher. I can write the things, which I cannot say in the classroom, here."* *"There is nothing special I shared with my teacher."*
	I write the things I can not understand in my journals.	8	53,3	6	50	
	I imagine the things our teacher expressed in the lesson while I write my journal.	-	-	1	8,3	
	We make ourselves known to our teacher and we know our teacher better.	-	-	1	8,3	
Yes, I shared.		12	80	10	83,3	
No, I did not share.		3	20	2	16,6	

Question 8. Do you want to continue logging in journals in science and technology and mathematics courses? Do you want to write journals in other courses?

According to the answers of the students have given, it seems that many of the students want to continue writing journals in the science-technology, math and some other courses. Percentage and frequency values for answers to the eighth question are presented in Table-8.

Table 8. Student opinions about eighth question and percentage-frequency values

Do you want to continue logging in journals in science and technology and mathematics courses? Do you want to write journals in other courses?

	Codes	6th grade f	6th grade %	7th grade f	7th grade %	Examples of student expressions
I prefer doing it for science-technology and math lessons, because…	To repeat the subjects that I could not understand	2	13,3	5	41,6	-*"…I have trouble with math and science. Therefore I would repeat more. I do not have any difficulty in other lessons. Therefore I would not."* -*"Yes, I do. For example I repeat the things I do not understand."* -*"I do in order to repeat the subjects."* -*"….because it makes me repeat. I would continue for Turkish, social and music."* -*"….since it is better. I do not want to write in other courses."*
	To keep the subjects in my mind	5	33,3	2	16,6	
	Because they are pretty.	5	33,3	2	16,6	
	To remember my childhood years	-	-	1	8,3	
	To have a better understanding of lessons.	4	26,6	2	16,6	
In other lessons	Yes, I do.	15	100	8	66,6	
	No, I do not.	-	-	4	33,3	

Conclusions and Implications

Results obtained from interview with the students indicate that writing journals is an entertaining activity. By the use of journals, communication increases increasing with the teacher because they can share the feelings and ideas smoothly. They contribute to repeating lessons, preparing exams, complying with the punctuation and spelling rules, taking care of the order of the writing and having permanent knowledge. In this scope journals can be used to help students express feelings and ideas, repeat the knowledge gained in the lessons and build up better student-teacher communication in lessons.

References

Çepni, S. (2005). Kuramdan uygulamaya fen ve teknoloji öğretimi. Ankara: Pegem A Yayıncılık.

Hargrove, T. & Nesbit, C. (2003). Science notebooks: Tools for increasing achievement across the curriculum. Retrieved December 15, 2007, from http://www.ericdigests.org/2004-4/notebooks.htm

Kaptan, F., (1998). Fen Bilgisi Öğretimi. Ankara: Anı Yayıncılık.

Korkmaz, H. (2004). Fen ve teknoloji eğitiminde alternatif değerlendirme yaklaşımları. Ankara: Yeryüzü Yayınevi.

Llewellyn, D. (2002). Inquire within implementing inquiry-based science standarts. California: Corwin Press.

MEB, (2006). İlköğretim fen ve teknoloji dersi öğretim programı. Talim Terbiye Kurulu Başkanlığı. Ankara: Devlet Kitapları Müdürlüğü. Retrieved December 15, 2007, from http://ttkb.meb.gov.tr/ogretmen/modules.php?name=Downloads&d_op=viewdownload&cid=74

Nesbit, C. R., Hargrove, T. Y., Harrelson, L. & Maxey, B. (2004). Implementing science notebooks in the primary grades. Science Activities: Classroom Projects and Curriculum Ideas, 40(4), 21-29.

Ruiz-Primo, M.A., Li, M., Ayala, C. & Shavelson, R. J. (1999, March). Student science journals and the evidence they provide: Classroom learning and opportunity to learn. Paper Presented at the NARST Annual Meeting, Boston, MA.

Ruiz-Primo, M.A., Li, M., Ayala, C. & Shavelson, R. J. (2004). Evaluating students' science notebooks as an assessment tool. International Journal of Science Education, 26 (12), 1477-1506.

Ruiz-Primo, M.A., Li, M. & Shavelson, R. J. (2002, April). Looking into students' science notebooks: What do teachers do with them? (CSE Tech. Rep. No. 562). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Shepardson, D. P. (1997). Of butterflies and beetles: First grader' ways of seeing and talking about insect life cycles. Journal of Research in Science Teaching, 34 (9), 873-889.

Shepardson, D. P. & Britsch, S.J. (2001). The role of children's journals in elementary school science activities. Journal of Research in Science Teaching, 38 (1), 43-69.

Tektaş-Hasanoğlu, A. (2002). Matematik günlüklerinin öğrencilerin matematik başarıları, matematiğe karşı olan tutumları ve matematik kaygıları üzerindeki etkileri. Unpublished master thesis, İstanbul: Boğaziçi University.

SCIENCE TEACHERS' COMPETENCIES ABOUT ALTERNATIVE ASSESSMENT METHODS

Havva Demirelli
Gazi University

Sedef Canbazoğlu
Aksaray University

Nusret Kavak
Gazi University

Nebahat Bekci
Gazi University

Abstract

Assessment is the process of identifying, gathering and interpreting information about students' knowledge and skills. Alternative assessment provides an alternative to paper-and-pencil testing. During preparing and using the alternative assessment methods, teachers need to have many competencies. This study is aimed to investigate the competencies of the science teachers at the stage of portfolio, rubric, concept map and structural communication grid methods. The data collection was gathered by two scales which were applied 98 science teachers in 2008-2009. Two scales consist of total 40 items concerning what should be done in the stage of preparing and using the mentioned methods. Results of the study indicated that the teachers follow the developments in the field of assessment and they have pointed out that they are in positive dealing on the alternative assessment methods and more than half of teachers didn't answered to the scale items concerning structural communication grid. When the structural communication grid method is investigated according to teachers' professional experience and number of students in their classroom, they rarely use this method. This research results will contribute to the field of teacher training and assessment.

Introduction

Assessment is the process of identifying, gathering and interpreting information about students' knowledge and skills (Popham, 2000). Alternative assessment provides an alternative to paper-and-pencil testing (Gronlund, 2006: p.223), and "performance assessment, portfolio assessment, peer assessment, self assessment, rubrics, concept maps, diagnostic tree, observation, structural communication grids" are several types of alternative assessment methods (Bol, et al., 2002; Johnstone, et al., 2000). Among these methods, portfolio is a systematic collection of students' works that are representative of student achievement at different points of learning process (Downing, 2005). Rubrics are used as the purpose of grading at most of the alternative methods of measuring and evaluating foremost portfolio. A rubric is a scoring guide which is used when measuring the level of performance or quality of a student product (Mertler, 2001), and rubric's scoring schemes are developed by teachers or other evaluators (Brookhart, 1999). Also structural communication grid is a flexible assessment method which can be used understanding relationships among concepts and facts (Johnstone & Ambusaıdı, 2001; Johnstone, et al., 2000). While applying this method, the teacher places the concepts, pictures, numbers, equations, definitions or formulas concerning the subject randomly. Giving the students different questions about the subject, they are requested to find out the convenient boxes for the answer of each question, and to line up these box numbers according to logical or functional order (Bahar et al., 2006). The other alternative assessment tool is concept map which is used as an assessment tool for student's understanding of scientific concepts (Robinson, 1999; Novak & Gowin,1984). During preparing and using the alternative assessment methods, the teachers need to have many competencies. This study is aimed to investigate the science teachers' competencies at the stage of portfolio, rubric, concept map and structural communication grid methods. In this context, we have compared teachers' frequencies of these methods according to their professional experience and number of students in the classroom.

Rationale

In recent years, as parallel to the constructivist approach, alternative assessment methods have been mostly used with the traditional methods. If the teachers don't have the enough knowledge about alternative methods, they will meet big problems while using these methods at assessment process (Hambleton & Murphy, 1992). Examining the competencies of teachers concerning portfolio, rubric, concept map and structural communication grid methods which are suggested to use especially curriculum will contribute the field of teacher training.

Methods

This descriptive research was performed in 2008-2009, and its sampling is constituted by 98 science teachers. The number of teachers across different professional experience is summarized in Table 1.

Table 1. The sample of the study according to teachers' professional experience

Professional Experience	Science Teachers
0 - 2 years	26
3 - 6 years	24
7 - 15 years	23
More than 16 years	25
Total	98

In the research, as the means of data gathering, the scales which are developed by researchers have been used. The items in the scales have been prepared by considering the opinions of experts concerning the studies carried out in literature about alternative and traditional assessment methods. The first scale was applied consists of 14 items which use a 5 point likert scale (1=strongly disagree, 5=strongly agree) to indicate the general opinion of teachers about assessment process. The second scale used in the study consists of two parts. In the first part, there are two questions regarding to the professional experiences of teachers and the average numbers of students in their classroom. In the second part consists of 26 items on a 3 point likert scale (1= never, 3=always) concerning what should be done in the stage of preparing and using portfolio (6 items), rubric (6 items), concept map (6 items) and structural communication grid methods (8 items).

Results

The Cronbach's alpha reliability coefficient of the scale developed in order to determine the general thoughts of teachers about the assessment process is 0.83, and the reliability coefficient of the scale concerning the stage of preparing and using the alternative assessment methods determined in the scope of research was calculated as 0.96.

The percentage of teachers who agree with the expressions given about the assessment process changes between 5.1 % and 56.1%, and the percentage of teachers who are strongly agree is between 20.4 % and 58.2 %. 46.9 % of the teachers have marked the expression *I agree* the item "I follow the recent developments about assessment and evaluation", and 33.7 % of them have marked the expression *strongly agree*. While 46.9 % of the teachers *strongly agree* with the item “I believe that alternative assessment methods make students much more active and eager” , 40.8 % have marked *agree,* 10.2 % *not certain* , 2 % *disagree* and there have been no teachers who have marked the expression *strongly disagree.* 51 % of the teachers have marked the expression *disagree* the item “I perform the assessment and evaluation at the end of term” I perform the assessment and evaluation at the end of term”, 32.7 % of them have marked the expression *strongly agree.* (Table 2)

Table 2. The percentage of teachers’ responses to first scale

	Percent Respondents (N= 98)				
Items	Strongly Agree	Agree	Not Certain	Disagree	Strongly Disagree
1. I have enough knowledge about the alternative assessment methods such as diagnostic tree and structural communication grid	20.4	50.0	15.3	8.2	3.1
2. I follow the recent developments about assessment and evaluation	33.7	46.9	13.3	5.1	0.0
3. I believe the necessarily of applying alternative assessment methods in science classroom	49.0	43.9	5.1	0.0	1.0
4. I think the alternative assessment methods affect the individual development of students in a positive way	58.2	30.6	8.2	2.0	1.0
5. I use the alternative assessment methods in order to determine what students know	33.7	55.1	8.2	2.0	1.0
6. I use the alternative assessment methods to acknowledge the students well	28.6	56.1	10.2	4.1	1.0
7. I perform the assessment and evaluation at end of term	3.1	5.1	8.2	51.0	32.7
8. I give information to students, families and colleagues about my assessment methods	25.5	56.1	10.2	7.1	0.0
9. I believe that alternative assessment methods allow students cooperate with each other	30.6	52.0	12.2	3.1	2.0
10. I believe that alternative assessment methods give students opportunity of show their abilities	39.8	48.0	9.2	1.0	1.0
11. I believe that alternative assessment methods make students much more active and eager	46.9	40.8	10.2	2.0	0.0
12. I also incorporate students into the assessment and evaluation process	33.7	48.0	12.2	6.1	0.0
13. My assessment activities are independent from each other	15.3	36.7	17.3	20.4	9.2
14. I prefer the methods which put the students to center in assessment process	36.7	49.0	10.2	4.1	0.0

The items of the second scale applied to the teachers are included by the expressions about the stage of preparing and using portfolio, rubric, concept map and structural communication grid methods. In this scale, concerning the method of portfolio, the teachers between 50 % and 71.4 % have expressed that they always do the things necessary to do at the stages of preparing and using the method. For instance, while 71.4 % of the teachers have marked the item “I explain students what the portfolio include” as always, 19.4 % of them as sometimes, and there have been no teachers who have marked the expression never. Over the scale items concerning portfolio, for the item “I interview with the parents to learn their opinion” 10.2 % of the teachers have marked the expression always, 59.2 % sometimes, and 22.4% never. (Table 3)

Table 3. The percentage and frequency of teacher responses: Portfolio

Items	Percent Respondents (N= 98)			
	Always	Sometimes	Never	No answer
1. I determine with the portfolio what behaviors being evaluated will be	50.0	41.8	1.0	7.1
2. I introduce portfolio to students	61.2	0.6	0.0	7.1
3. I explain students what the portfolio include	71.4	19.4	0.0	7.1
4. I inform students about how portfolio will be prepared and when it will be ready	65.3	26.5	1.0	7.1
5. I interview with the parents to learn their opinion	10.2	59.2	22.4	7.1
6. I improve some scales according to workouts taken place in portfolio	43.9	41.8	7.1	7.1

The answers that the teachers have responded to the scale items concerning the stages of preparing and using rubric and concept map methods have indicated the distribution as always between 50 % and 74.5 %, as sometimes between 16.3 % and 37.8 %, and as never between 1 % and 6.1 %. For instance, concerning the rubrics, for the item "I determine the goals of rubric", 63.3 % of the teachers have given the answer *always.* Concerning the concept map, 73.5 % of the teachers have given the answer *always* for the item "I determine the key concepts which are necessary for being understood a subject". (Table 4)

Table 4. The percentage and frequency of teacher responses: Concept map

Items	Percent Respondents (N= 98)			
	Always	Sometimes	Never	No answer
1. I determine the key concepts which are necessary for being understood a subject	73.5	21.4	0.0	5.1
2. I give some clues to students in order to choose right concepts	64.3	30.6	0.0	5.1
3. I order concepts from general to specific one	62.2	30.6	2.0	5.1
4. I determine the relations between concepts by means of arrows	74.5	20.4	0.0	5.1
5. I make cross-connection between concepts	54.1	37.8	3.1	5.1
6. I name the cross and inter connection I have determined	55.1	32.7	6.1	6.1

55 % of the teachers haven't answered to the scale item concerning the structural communication grid method. When it is investigated what occasions the teachers use this method according to the professional experience and the number of the students in the classrooms, it is understood that they have marked the expression *never* in the scale. (Table 5) Hence, it is determined that structural communication grid method is used less than the other methods.

Table 5. The percentage and frequency of teacher responses: Structural communication grid

Items	Percent Respondents (N= 98)			
	Always	Sometimes	Never	No answer
1. I make a list of concepts or samples related to topic	31.6	13.3	0.0	55.1
2. I prepare a table depending on the numbers of concepts and samples	25.5	18.4	1.0	55.1
3. I place the concepts or samples to the table randomly	22.4	13.3	9.2	55.1
4. I give a name as A,B,C or number as 1,2,3 to each cell	24.5	18.4	2.0	55.1
5. I prepare questions in such a way that none of the concepts or the samples in the table remain uncovered	26.5	15.3	3.1	55.1
6. I determine the number of boxes (C2) which can be right answer of each question and the number of boxes (C4) which is accepted as wrong	17.3	21.4	6.1	55.1
7. I determine the number of each right answers of student (C1) and the number of his wrong answers(C3)	20.4	18.4	6.1	55.1
8. I use the formula of C1/C2- C3/C4 to give points when the student finds the right box for the answer of each question)	15.3	22.4	7.1	55.1

Conclusions and Implications

According to the results of the research, the teachers have stated that they follow the developments in the field of assessment, and they have pointed out that they are in positive dealing on the alternative assessment methods. For the details concerning the stages of preparing and using portfolio, rubric, and concept map searched in the scope of study, more than half of the teachers have stated that they always carry out, but more than half of them haven't answered to the scale items concerning structural communication grid. When the results of teacher's using the structural communication grid method is investigated according to professional experience and number of students, it is understood that they have indicated that they rarely use this method. In the lights of findings of the study, it is understood that teachers need education on using and preparing alternative assessment methods. It is stated that teachers have not been informed adequately about alternative assessment methods that are suggested to be used in science curriculum. These methods such as structural communication grid, diagnostic tree have not been understood by teachers (Gözütok et al., 2005). It can be thought that the proficiency of teachers concerning this method is limited. If the literature is examined, it can be seen that structural communication grid method is mainly used for determining the students' misconceptions (Bahar et al., 2002). However, in this study, teachers' proficiency has been investigated on the use of structural communication grid during assessment process.

Although the alternative assessment methods take up more time according to the traditional methods at the process, the science teachers in the scope of research have stated that they often use the alternative assessment methods. Furthermore, any relation hasn't been seen between what occasions the teachers use the alternative methods and class sizes. In the similar way, a certain relation hasn't revealed between what occasions the teachers use the methods and professional experiences.

In this study, no differences have been found between teachers' applying to these alternative assessment methods. But Duban and Küçükyılmaz (2008) found that teachers use the portfolio method more than rubric and concept map.

In Turkey, science and technology curriculum (Ministry of National Education [MEB], 2005, 2006a) recommends using alternative assessment methods in student textbooks and in workbooks. Generic teacher competencies also were introduced in 2006. These competencies, which emphasize the alternative assessment methods, will prove very useful in terms of identifying task definitions of teachers. *Monitoring and evaluation of learning and development* is main competency and *identifies alternative testing tools for a comprehensive assessment (portfolios, concept maps, observations, interviews and etc) is* the sub-competency of generic teacher competencies (Ministry of National Education [MEB], 2006b). In order to use alternative assessment methods effectively, teachers must have some competencies about these methods.

In order to have adequate knowledge about alternative assessment methods; teachers should participate in-service training courses more frequently. Some of researchers gave in-service training course about structural communication grid, diagnostic tree and portfolio and they have resulted that in-service training courses have positive effect upon teachers' abilities and using the alternative assessment methods (Çoruhlu et al., 2008). What's more, during the courses alternative assessment methods must be emphasized and the number of courses related to these methods should be increased.

In the light of the findings obtained from this research, doing interviews with teachers who have different professional experience on alternative assessment methods and examining the documents (exam papers, worksheets) that are used by teachers during assessment process will contribute to the science education especially field assessment and evaluation.

References

Bahar, M., Nartgün, Z., Durmuş, S., & Bıçak, B. (2006) .*Geleneksel-alternatif ölçme ve değerlendirme teknikleri, öğretmen el kitabı* [Traditional-alternative assessment and evaluation methods, teacher handbook] . Ankara: Pegem A Yayıncılık.

Bahar, M., Öztürk, E., & Ateş, S. (2002). *Yapılandırılmış grid metodu ile lise öğrencilerinin Newton'un hareket yasası, iş, güç ve enerji konusundaki anlama düzeyleri ve hatalı kavramlarının tespiti* [Determining misconceptions and understanding level of high school students about Newton's laws of motion, work, power and energy topic with structural communication grid]. V. Ulusal Fen Bilimleri ve Matematik Eğitimi Kongresi, ODTÜ, Ankara.

Bol, L., Ross, S.M., Nunnery, J.A., & Alberg, M.S. (2002). A comparison of teachers' assessment practices in school restructuring models by year of implementation. *Journal of Education for Students Placed at Risk, 7*(4),407-423.

Brookhart, S. M. (1999). The art and science of classroom assessment: The missing part of pedagogy. *ASHE-ERIC Higher Education Report, 27* (1).

Çoruhlu,T.Ş., Er Nas, S., & Çepni, S. (2008). Reflection of an In-service Education Course Programme for Science and Technology Teachers about Alternative Measurement and Assessment Techniques: Trabzon Sample. *Necatibey Faculty of Education Electronic Journal of Science and Mathematics Education, 2*(2), 1-22.

Downing,J. (2005). Teaching literacy to students with significant disabilities: Strategies for the K-12 inclusive classroom. Thousand Oak, CA: Corwin Press.

Duban, N. & Küçükyılmaz, E.A. (2008). Primary education pre-service teacher's opinions regarding to the use of alternative measurement-evaluation methods and techniques in practice schools. Elementary Education Online, 7(3), 769-784. Retrieved January 10, 2009 from http://ilkogretim-online.org.tr/.

Gronlund, N. E. (2006). *Assessment of student achievement* (8th edition). Boston: Pearson Publishers.

Gözütok, D., Akgün, Ö. E., & Karacaoğlu, C., (2005). *Yeni ilköğretim programlarının uygulanmasına öğretmenlerin hazırlanması* [Preparation of teachers in the implementation of new elementary curriculum]. Eğitimde Yansımalar: VIII. Yeni İlköğretim Programlarını Değerlendirme Sempozyumu, Kayseri.

Hambleton, R. K., & Murphy, E. (1992). A psychometric perspective on authentic measurement. *Applied Measurement in Education, 5*(1), 1–16.

Johnstone, A.H., & Ambusaıdı, A. (2001). Fixed-response questions with a difference. *Chemistry Education: Research and Practice in Europe, 2*(3), 313-327.

Johnstone, A.H., Bahar, M., & Hansell, M.H. (2000). Structural communication grids: A valuable assessment and diagnostic tool for science teachers. *Journal of Biological Education, 34*(2), 87-89.

Milli Eğitim Bakanlığı [Ministry of National Education]. (2005). *İlköğretim fen ve teknoloji dersi (4.-5.sınıflar) öğretim programı.* [Elementary science and technology lesson (4th-5th grade) instruction program]. Ankara: Author.

Milli Eğitim Bakanlığı [Ministry of National Education]. (2006a). *İlköğretim fen ve teknoloji dersi (6.-8.sınıflar) öğretim programı.* [Elementary science and technology lesson (6th-8th grade) instruction program]. Ankara: Author.

Milli Eğitim Bakanlığı [Ministry of National Education]. (2006b). Support to basic education project:: Generic teacher competencies. Ankara: Author.

Mertler, C. A. (2001). Designing scoring rubrics for your classroom. *Practical Assessment, Research & Evaluation, 7*(25).

Novak, J.D., & Gowin, D.B (1984). *Learning how to learn.* New York: Cambridge University Press

Popham, J. W. (2000). *Modern educational measurement.* Needham: Allyn & Bacon

Robinson,W.R. (1999). A view from the science education research literature: Concept map assessment of classroom learning. *Journal of Chemical Education,* 76 (8), 2-3.

Stone, S. (2004). *Multi age.* Good year books, Glenview, III.

MEASUREMENT OF CONTENT RELATED ASPECTS OF SCIENCE COMPETENCE

Alexander Kauertz
University of Education of Weingarten

Hans E. Fischer
University of Duisburg-Essen

Peter Labudde
University of Applied Sciences Northwestern Switzerland

Hendrik Härtig
University of Duisburg-Essen

Mathias Ropohl
University of Duisburg-Essen

Maik Walpuski
University of Osnabrueck

Elke Sumfleth
University of Duisburg-Essen

Abstract

Measurement of competencies in science is an important tool to inform school administrators and educators about the status of the school system and a realistic view on mean student abilities. As consequences there might be changes in the curriculum or the standardization of the national education standards. The content used in the items for the assessment defines the subject which is measured and influences the item difficulty. Therefore the role of content for the assessment needs to be discussed. In this chapter authors from two large assessment projects in Germany and Switzerland present the respective assessment and how content validity is ensured and the content's influence on difficulty is investigated. The first two authors briefly introduce to the assessments' frameworks. The following two authors report about research studies conducted within the second assessment project.

Introduction

Measuring science competence in large-scale-assessments requires a choice of content for the tasks used in the paper-and-pencil assessments. The role of content for assessing competence is illustrated using the Swiss and German educational standards and the corresponding assessments that have been developed. While content is used as one dimension of the Swiss competence model, in Germany content is instead considered implicitly in the different areas of competence. In both cases, the effect on the difficulty of tasks is discussed. In a different approach, the role of content can be analysed in comparison to the content used in classrooms, leading to the question of content validity of competence tests. A typical technique in making tasks in the assessment independent of the curriculum is to provide all necessary information in the task. This causes an effect on the validity of the test's interpretation which should be critically discussed. Because competence is a complex concept, the role of content is related to many different facets of competence assessment.

Ever since McClelland (1973) defined competence in contrast to intelligence, there has been an ongoing debate on how competence can be described as well as what contributes to it. A widely accepted interpretation is used in the large-scale-assessment of the Programme for International Student Assessment (PISA, OECD, 1999). This notion of competence has been adapted for defining standards in different countries. These standards come along with large-scale-assessments to evaluate the school system.

For those assessments, competence is operationalized in competence models (Schecker & Parchmann, 2007). Competence models often use several dimensions. One dimension usually describes the content used in the tasks. A certain competence therefore combines a cognitive operation, for example, with certain content on a specific proficiency level. Therefore, competence describes more than just having knowledge in a domain.

In Switzerland, a competency model, standards and an assessment have been developed by a research group (2005-2008). In the first presentation, we will present results from this Swiss project that allow the assessment of the quality of the used competence model and will also propose standards to appropriate policy makers. The tests and results are focused both on the domain of content as on different competencies.

In Germany, standards for science have been defined by The Standing Conference of the Ministers of Education and Cultural Affairs of the States (Bundesländer) in the Federal Republic of Germany (KMK, 2005a, 2005b). A project has been set up to develop an assessment to evaluate these standards (Walpuski, Kampa, Kauertz & Wellnitz, 2008). Unlike in the Swiss model, the content is not a dimension in and of itself but rather dealing with scientific content according to four basic concepts is one area of competence. The second contributor will show how the difficulty of tasks is varied by complexity and cognitive processes in this field.

Although both approaches use scientific content in different ways to define scientific competencies, they use content in tasks for large-scale assessments. That way, both approaches have to show that they are valid for the respective curricula and regarding the taught content in classrooms. Since validity is known as a central problem in large-scale assessment, the third author is presenting a method to analyse how validly a task represents the content from the curriculum.

One appropriate method to control the influence of students' content knowledge on the measurement is to provide all necessary knowledge in the tasks. This method is used in the German assessment. In the final contribution the last author systematically investigates the effect of such presented information in tasks as part of the validity of the competence model.

All contributions to this symposium investigate how scientific content and scientific competence are related. In each contribution, the role of content is viewed differently, and we thus can illustrate in a wider range the effect of content on competence and discuss the validity of competence models and tests.

COMPETENCIES IN SCIENCE: A LARGE-SCALE ASSESSMENT IN SWITZERLAND IN GRADES 6 AND 9

Peter Labudde
University of Applied Sciences Northwestern Switzerland

In several European countries competence models and standards in science education have been (re)defined recently or are being defined (Waddington et al. 2007), in many cases initiated by the Programme for International Student Assessment (PISA). In order to evaluate the models, the standards, and/or the students, science educators have developed and accomplished large-scale assessments, e.g. in Switzerland a broad consortium of science educators and teachers carried this out on behalf of the policy makers (2005-2008).

What are the main competencies in science instruction that students have to achieve? What could and should be the level of their performance? In Switzerland a three-dimensional competency model has been developed (Labudde 2008). It includes competencies (1st axis), domains of contents (2nd), and levels of understanding (3rd). The first dimension comprises eight competencies, e.g. "exploiting information sources", "organising, structuring and modelling", "assessing and judging", and "asking questions and investigating". Several tests have been accomplished in order to validate the model and to get an empirical basis for the proposition of future basic standards for the end of grades 2, 6, and 9.

One of the tests was a large-scale evaluation at the end of grade 6 (N=4124, representative sample of the French and German speaking part of Switzerland; 45 PISA-like situations with totally 229 paper-and-pencil-items) and grade 9 (N=3888; 45 situations with 262 p&p-items). The items included multiple choice and short answer items. Each item was assigned to one of the first three competencies mentioned above and to one of the domains of contents. The difficulty of the items and the performance of an individual have been determined by means of a Raschanalysis.

In general, the items have been proven to be more difficult for the students than expected by science educators and teachers. As a consequence this has led to cautious and subtle propositions of basic standards in science instruction in Switzerland. Other test results are performance differences between boys and girls and between the German and French speaking parts of Switzerland. Further findings are the high correlations (0.72 - 0.94) between different competencies leading to the question, how far a definition and separation of different competencies in science are possible at all. In our talk we will present a detailed description of the competency model, give some typical examples of test items, and discuss the results from different perspectives: science education, daily science instruction, teacher professional development, and politics.

The research project contributes to the discussion and definition of competencies in science instructions, the validation of competency models, and the evaluation of students' performances in large-scale assessments. It yields also paradigmatic examples of good practice, both for instruction/learning and for evaluation.

EVALUATING THE GERMAN NATIONAL STANDARDS FOR THE SCIENCE SUBJECTS– THE COMPETENCE AREA "APPLICATION OF CONTENT KNOWLEDGE"

Alexander Kauertz
University of Education of Weingarten

Hans E. Fischer
University of Duisburg-Essen

Like many countries before, Germany has changed its educational system from an input-oriented paradigm to an output-oriented one by deciding on educational standards. To evaluate the educational system, a test has to be developed to adjust the standards to the students' abilities. This test can later be used to evaluate the educational system. Unlike in other countries, the German test is based upon a previously defined and empirically researched competence model. That way the results can be used to give insight into the competence structure of German students. Experienced teachers are asked to develop the tasks for the test, supported by science education researchers and psychologists who provide guidelines for task development. These guidelines help teachers to translate the competence model into tasks.

The implemented competence model is an elaborated version of the competence model that is presented in the German standards for sciences at lower secondary school level (KMK, 2005a, 2005b) and has been specified for biology, chemistry and physics. According to this model, competence is described by three dimensions: complexity, cognitive processes and competence area. As for complexity, different levels of competence are defined based on the content structure of the task (c.f. Kauertz & Fischer, 2006). Cognitive processes describe the way students have to deal with the information in the task. Finally, four areas of competence are distinguished: dealing with scientific knowledge, working scientifically, communicating, and judging. In addition, the content area of the expected scientific knowledge at the end of Grade 10 is structured using four basic concepts: energy, matter, interaction and system. Accordingly, each task of the test can be characterized by one basic concept, a specific complexity level and requires a certain cognitive processes for solving. One main purpose of the presented study is to show that even low-performing students are able to solve some of the tasks, and thus the test is able differentiate competence even at a low-performer level. This seems to be of crucial importance in order to find ways to foster those students rather then labeling them as "not good enough".

For the competence area "dealing with scientific knowledge," 142 paper-and-pencil-tasks have been developed by teachers for all combinations of basic concepts, complexity and cognitive processes. N = 296 students from Grade 10 worked on the tasks. For analysis, the number of correct solutions per task is counted and the standard deviation of each item is calculated. For selected tasks, the students are asked to rate on 4-point Likert-scales how familiar they feel with the content of the task and how interesting the task is for them.

The solution frequency of the tasks varies between zero solutions and 90% correct solutions (mean solution frequency is 56%). The mean standard deviation of all items is SD=.44, the maximum is Max=.51. The familiarity of the content is slightly under the average (M=2.44, SD=.53), the interest in the tasks is even smaller (M=2.13, SD=.51). The correlation between familiarity and interest is r=.64 ($p<.001$). For most tasks, no correlation between familiarity or interest and correct answer is found.

The full range of students' competence is covered by the tasks, as the range in number of solutions indicates. The standard deviation shows that all tasks can differentiate between different students' competence. On average, the content of the tasks is familiar to the students, but the students are instead uninterested in the topic of the tasks. The students' success in most tasks is independent from their familiarity with the content or their interest in the content. This provides a first hint that the tasks successfully measure the competence to handle the scientific content rather than knowing or being interested in the content itself.

ANALYSIS OF THE INFLUENCE OF GIVEN CONTENT INFORMATION ON THE ACHIEVEMENT IN A COMPETENCY TEST

Mathias Ropohl
University of Duisburg-Essen

Maik Walpuski
University of Osnabrueck

Elke Sumfleth
University of Duisburg-Essen

Abstract

Results of international comparative large-scale assessments like PISA (Baumert et al., 2001) revealed serious flaws in the German school system. As a response, students' intended learning outcomes have been defined by national educational standards, amongst others in chemistry (KMK, 2005). They describe competences students should have acquired when they receive their general educational school-leaving certificate after grade 10. But, the competences are presented without a description of specific contents. Therefore, the issue of an evaluation of the standards with a paper-pencil-test have to be defined by a content analysis of chemistry curricula. There are just a few overlaps with regard to common contents between the curricula from the different states. Thus, the tasks for the evaluation of the national educational standards provide content information in the item stem. The students have to use this information to work on the tasks. In this PhD-study the influence of the presentation of content information on the item difficulty is analysed by a second task type. In addition, the predictability of the item difficulty by a normative structure model of competence is specifically investigated.

Introduction

For the evaluation of normatively determined competences appropriate tasks have to be developed. To link the standards with task formulations Klieme et al. (2003, p. 16) require the development of a normative structure model of competence. By means of the model competences can be operationalised based on certain characteristics.

Science education researchers have developed a normative structure model of competence for all science subjects (Walpuski, Kampa, Kauertz, & Wellnitz, 2008). In accordance with the standards the model differentiates between various areas of scientific competence: "application of content knowledge", "acquirement of knowledge", "communication" and "evaluation and judgment". Furthermore, the model contains two additional dimensions: "complexity" and "cognitive processes" (see figure 1). The present study focuses on the application of chemical content knowledge combined with the other dimensions.

The development of the dimension "complexity" is based on precedent studies published by Commons et al. (2007) and Kauertz (2008). In view of the area of competence "application of content knowledge" complexity means the complexity of content. This dimension is supposed to contain hierarchical levels. Cognitive processes are defined as processes test persons have to use when they solve tasks.

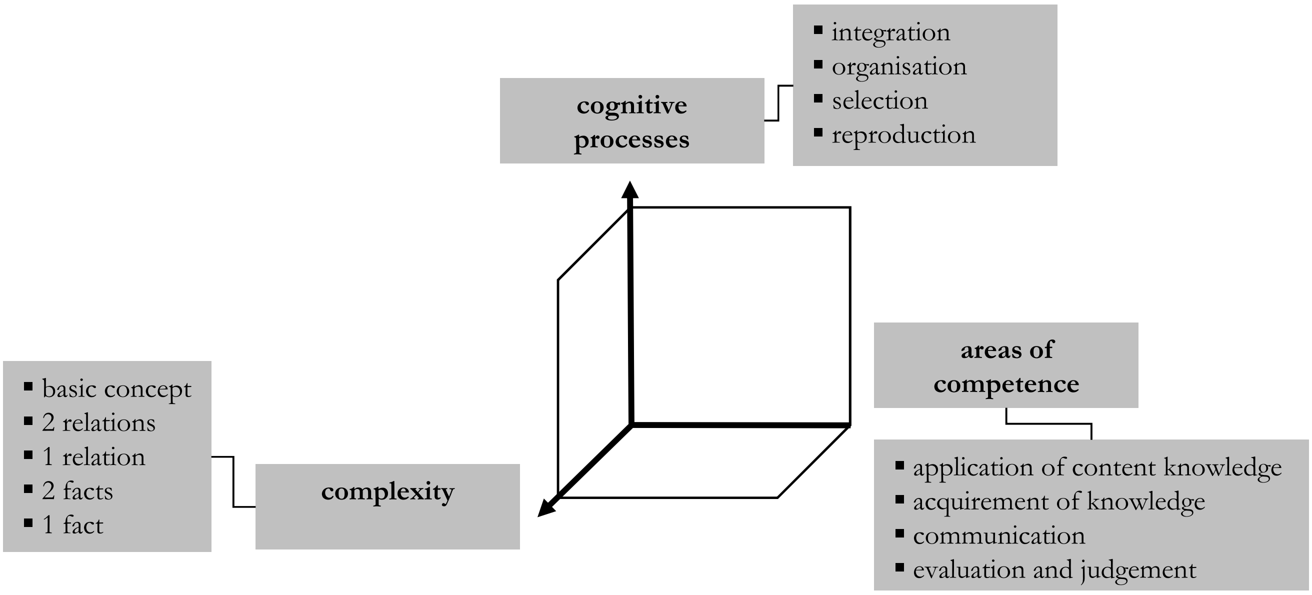


Figure 1: structure model of competence (Walpuski, Kampa, Kauertz, & Wellnitz, 2008)

Rationale

In connection with the evaluation of the national educational standards content information needed for the solution of the tasks will be presented in the item stem due to different curricula for chemistry in all 16 German states. Each state has curricula for every subject and every school type. Thus the contents in chemistry can vary from state to state and it is difficult to determine the topics. The overlap of the actually taught content is quite small.

The content information is given in a short text followed by assignments (task type A). This procedure is necessary to ensure that the tasks don't aim at the interrogation of content knowledge students have but at the assessment of the application of chemical content knowledge. By giving all relevant information it is intended to minimise the influence of the previous knowledge on the observed achievement.

The main goal of the PhD-study is an analysis of the influence of the previous knowledge on item difficulty. For this purpose, a second type of standard-based tasks has been developed (task type B). Its distinctive characteristic is the absence of content specific information from task type A. Thus, students have to depend on their previous knowledge.

Table 1. Task type A vs. task type B.

Task type A Chemical content knowledge **provided**	*Task type B* Chemical content knowledge **presupposed**
Item stem The item stem **includes relevant information** to answer the question. Thereby, the influence of the previous knowledge on the item difficulty shall be minimised.	*Item stem* The item stem **does not include relevant information**. The students have to solve the tasks basing on their previous knowledge.
Item question or task instruction	*Item* identical to type A
Options 1 correct choice 3 distractors	*Options* identical to type A

In this context, the following research questions are posed:

1) What influence on achievement does the presentation of necessary content information have in comparison to items without this information?

2) How does the empirical data of a Rasch analysis fit to the assumed model of competence – especially for the area of competence "application of chemical content knowledge" and the subarea "chemical reactions"?

Methods

According to the structure model of competence standard based test items are developed for the above named area of competence. The development process is standardised by a construction guideline (Köller et al., 2008) which contains general rules concerning the test item design as well as information about the implementation of complexity and cognitive processes in a test item. Thereby, every test item is referred to a combination of complexity and cognitive process. Power analysis advises a sample of 160 items to differentiate between the two task types. 80 tasks of type A and 80 of type B were developed and allocated to 16 booklets in a multi-matrix-design. Each booklet contains 20 items, 10 of each type.

The model-fit of the developed test items was investigated with Rasch analysis (Kauertz & Fischer, 2006). It is expected that item difficulty increases with complexity of content and with the demand of the cognitive processes needed to answer the item.

Furthermore, the variables last mark in chemistry and cognitive skills (Heller & Perleth, 2000) were measured. It is expected that the probability of the correct answer correlates to some extent with reading comprehension. Therefore, the study contains a third test to enable control of this influence.

Results

During the first stage of the project, the content for the test items was validated by an evaluation of current curricula with respect to technical terms. The terms were listed in tables assigned to different contents of chemistry and to different stages of education. Afterwards they were analysed by a rating of experts. The inter-rater reliability lies in a very good interval (Kappa=.87 and Kappa=.97) (Wirtz & Caspar, 2002). On the basis of this analysis, a high content validity of the tasks can be ensured because the analysed technical terms are used to develop the standard-based tasks.

In the second phase 1.500 students of 10^{th} grade from Northrhine-Westphalia participated in the assessment. The sample size varies between the different tests, because some students were absent at one of the two test days. Following there is given a survey of first findings from the tests.

The 160 test items were done by 1.365 students. Due to the Multi-Matrix-Design there are about 170 student responses per item. The model fit of the test items was proved with the weighted mean square value (MNSQ) and the t-value belonging to it. In this analysis the characteristic values fit the usual criteria of quality (Wilson, 2005, p. 129; Wu, Adams, Wilson, & Haldane, 2007, p. 23). Therefore, no items have to be deleted from the further analysis. The estimated item difficulties and person parameters from this Rasch analysis are the basis for the following evaluation.

First, the analyses of the item difficulty focuses on the adapted dimension „complexity of the content", which Kauertz (2008) identified as difficulty causing. At the beginning, the two task types were considered separately. For the evaluation of the differences between the mean values of the stages of complexity t-tests were computed. The item difficulty of task type A differs between most of the stages of complexity with one exception significantly. In contrast to task type A for task type B significant differences could not be determined among the stages. Afterwards, the differences of the mean values were regarded between the task types on one level of complexity. The differences between the task types decrease with increasing complexity. At the moment further analysis of the test results are realised to measure the influence of the ascertained companion variables on the achievement in the test of application of chemical content knowledge.

Conclusions and Implications

Furthermore, 24 test items are employed in a state-wide assessment of the evaluation of the national educational standards. In spring 2010 further data are there, which can be directly compared to the results of this study.

Lately, the study contributes to the evaluation and further development of the educational standards. Standard-based test items will be developed to assess students' knowledge in chemistry at the end of 10th grade.

In addition, this study contributes to the validation of the assumed competence model, especially with regard to cognitive processes. At the moment, no current cognitive process taxonomy seems validated by adequate theoretical development and research (Haladyna, 2004).

CONTENT VALIDITY IN LARGE-SCALE-ASSESSMENTS

Hendrik Härtig
University of Duisburg-Essen

Alexander Kauertz
University of Education of Weingarten

Hans E. Fischer
University of Duisburg-Essen

Abstract

Supported by the Standing Conference of the Ministers of Education and Cultural Affairs (KMK) of Germany, a cooperative project is currently developing a nationwide assessment system for junior high level standards in Biology, Chemistry and Physics. An internationally well known problem concerning large-scale assessments is content validity (Lissitz & Samuelsen, 2006). Content validity focuses on the correspondence between what should be (or is) taught in schools and content covered by an assessment tool. To ensure content validity of the mentioned assessments, content and its structure have to be reconstructed in a way that enables quantitative analyses. In this study, a method is developed and validated to ensure content validity of the assessments. Since teachers often use textbooks in preparing their lessons, textbooks can tell us what is taught in schools in a better way than curricula (Kirk, Matthews, & Kurtts, 2001). Therefore the content of widely used textbooks is coded using concept maps. Concept maps can be analyzed with different parameters to develop a new approach for item development. The satisfying results of the mapping method will be presented.

Introduction

In 2005 the Standing Conference of the Ministers of Education and Cultural Affairs (KMK) of Germany passed national education standards (NES) for secondary school exams in Languages, Mathematics, Biology, Chemistry and Physics (KMK, 2005). Students are expected to meet these benchmarks, and therefore NES are achievement standards for the junior high level up to grade 10 (Schecker & Parchmann, 2007). In contrast to standards for example in the USA the NES do not include school-system or teaching standards. NES follow Weinert (2001) and focus on competencies. So they are performance standards and explicitly not content standards. One competency for example is: "Students are able to recall physics concepts, measurement methods, variables and laws" (KMK, 2005, p. 11). As one can see, content knowledge is mentioned on the whole but the content is not explicitly described.

Due to the federal structure, each state develops its own curricula to meet these performance standards (State Performance Standards / SPS). Students' performance will be measured by centrally administered nationwide assessments for schools and students. Therefore the states' curricula, as well as what is taught, must correspond to NES. Unfortunately, neither NES nor SPS include content standards. As a result, teachers have considerable liberty with regard to the content they teach.

NES have to be assessed if they shall be implemented (Schecker & Parchmann, 2007). The development of national assessments is being accomplished through cooperation with science educators (Biology, Chemistry and Physics) and psychometricians. To ensure the quality of the test instrument, many types of validity analyses will be performed both nationally and in individual states. Test items always have to deal with content if they do not assess general abilities. As Lissitz & Samuelsen (2006) note, content validity was neglected in many large-scale assessments. Riffert (2005) points out that in international assessments, curricular validity was not in the focus due to differences between the curricula of the participating countries. Because of their federal structures, e.g. Canada, USA, and Switzerland have similar difficulties. Due to increasing orientation towards competence, NES and SPS include less and less detailed content description. In particular, the overall content structure cannot be found in German curricula. So federal structure encourages the use of nationwide standards and one referring test instrument, but regarding curricula it causes huge validity problems.

Rationale

Textbooks typically include the entire domain of knowledge for a particular subject area. Furthermore, authors of textbooks have to orientate the subject material towards curricula because in German textbooks may only be used after passing an accreditation. Additionally Shavelson (1971) was able to show a strong relationship between a presented textbook structure and the newly built cognitive structure of students in Physics. Therefore, curricular validity typically coincides with content validity. Assuming that teachers in many countries often use textbooks to prepare their lessons (Merzyn, 1994; Kirk, Matthews, & Kurtts, 2001), students are actually faced with textbook content which corresponds with the respective curriculum.

The core idea in this study is to develop a method for documenting the content and its overall structure out of textbooks similar to an approach of Vlug (1997) who analyzed textbooks to ensure the curricular validity of an assessment tool. The structure of the content is important to establish a hierarchy of importance among content topics. According to Chiappetta, Fillman, & Sethna (1991) or Wu, Lee, & Lai (2004) concept maps are appropriate methods to represent the content of textbooks and its structure. Moreover the analysis of concept maps leads to a differentiation between more or less central elements. Neumann, Trendel & Fischer (2006) report good experiences with analyzing curricula with concept maps, but they fail to delineate clear and sharp guidelines on how to extract concepts and propositions from the text. As all the mentioned mapping studies state a guideline is the only way to a reliable and valid coding process. So the challenges are to choose what text to extract extract this text following certain rules, place it in some sort of format on a concept map, and validating the maps as measure of content validity for the NES test items. These concept maps are a measurable way of presenting the content structure for teaching physics and will lead to a new approach for developing content valid assessments. To reach this goal, textbooks from different German states must be analyzed, reconstructed and compared accordingly.

Methods

The first aim is to develop a reliable method coding textbooks into concept maps. Additionally the maps shall be validated as predictor for what is taught in physics lessons and last be analyzed whether they might predict the content validity of NES test items. The procedure is subdivided into six steps:

Step 1: Choosing the texts to analyze

In Germany each publisher has different editions for most of the states and within the states again different editions for each school type (there are three school types which are differentiated by the proficiency level). Due to the large number of different textbooks a sample has to be chosen by interviewing the publishers. They reported that within one publishing company the different editions are all based on one book. To meet the curricula of the different states chapters are added or taken out. The publishers agreed upon the states with the biggest difference in the covered content. Because the textbooks differ a lot between different publishers the textbook sample should cover the four biggest publishing houses and the states with the biggest content differences. Six books were chosen; these books are often used and representative for the widest variance of content as we know from interviews with the publishers. All books include the grades seven to ten and they also cover two different complexity levels.

To analyze the validity of the sampling, the contents of the six chosen textbooks were compared with four randomly selected books. Therefore a list of topic was collected. Each topic should take one two three pages in the textbooks (e. g. "Ohm's Law" or "Newton's First Law"). The two samples (six concept mapped and four randomly selected textbooks) show a high correlation (N(topics)=288; Pearson's r=.663; p<.001) concerning the content they include.

Step 2: Separating scientific concepts from non-scientific concepts

A guideline for textbook analysis is necessary to guarantee a reliable, valid and objective procedure to transfer content of textbooks into concept maps. As concept maps should only contain scientific concepts knots, separating scientific concepts from non-scientific concepts was a major problem. Therefore, a definition of scientific concept was created, illustrated by a list of concepts and non-concepts: A noun (perhaps restricted through an adjective) or its pronoun is a scientific concept if it is well-defined in physics and used in a physics context. The decision is made upon university physics content knowledge. Two constraints are introduced: Measurement equipment and experiments shall only be coded if they are used to measure a physics variable explicitly or present the relationship of two variables. Additionally technical equipment is only taken as scientific concept if its meaning is not important in students' all-day life. Based on the definition and the list, six coders collected all concepts and a lot of non-concepts from various textbooks in a database as basis for the analysis. This database is extended during transferring the content of textbooks into concept maps if a new scientific concept is identified.

Step 3: Transferring the content into concept maps

To ensure a reliable coding process a guideline for transferring the content of the textbooks into concept maps was developed. The guidelines help to make uniform decisions; six coders were trained in identifying scientific concepts and in using the guideline to create concept maps. The training included how to use the mapping software Cmap tools© and to make similar decisions during drawing the concept maps. Using the list of concepts scientific concepts taught at school are identified and used as knots in the concept maps. Those sentences that describe or connect one or more concepts are extracted as linking words / sentences in the maps using Cmap tools©. One linkage can be connected to one knot or it connects two or more knots. An example is shown in figure 1.The concept maps can be exported into a content database which will be fundamental for developing content valid test items. If it is possible to generalize the resulting database, the map can be checked by determining the overlap of concepts from currently un-coded textbooks.

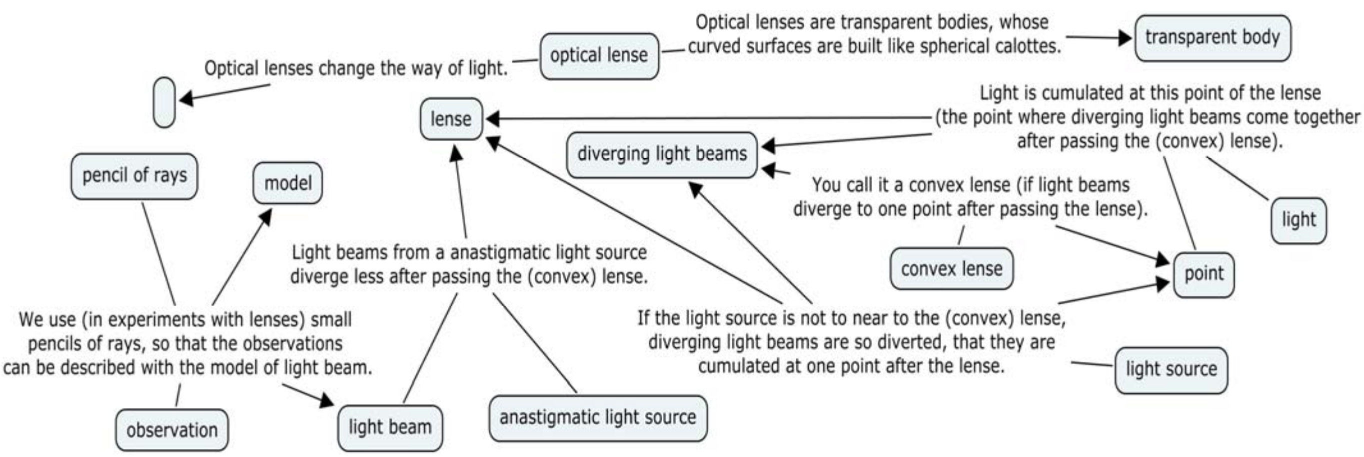


Figure 1: Exemplary Map

Step 4: Analyzing the quality of the method

The goal of the study presented is to develop a reliable and valid method for transferring textbooks content into concept maps. To analyze reliability, validity and objectivity of the applied concept mapping, different methods are used. Concerning step 2 firstly the coders have to rate a list of nouns - whether they are scientific concepts or not based on the described rules. Their agreement is measured in relation to an expert rating of 2000 nouns. For analyzing the quality of the third step, intercoder-reliability is analyzed as all coders analyze the same text. This text was taken from at least two different chapters from two different textbooks. If the coding process turns out to be reliable external validity must be shown. Therefore a comparison of the maps with videotaped lessons can show the relationship between the textbooks content and what is really taught in schools sounds suitable.

Step 5: Analyzing the concept maps and the NES items

To make it possible to compare the NES items with the concept maps, analyses of the maps are needed. Shavelson (1971) analyzed texts using the centrality of the concepts. He produced robust results by using this as indicator for students' understanding of the concepts provided in the texts. Therefore in the presented study all concepts in the maps were analyzed concerning their centrality; last each concept is described by one value. The NES items use physics terms, these terms can now be described by the centrality-value. All concepts that are used in one item have a specified centrality-value and as a first approach the mean of all values within one item shall be used as a measure of the validity of this item (validity-value for each item).

Step 6: Expert rating on the NES items and comparison to the concept maps

An expert rating on the content validity of fifteen of the NES items (which are representative for the test) was performed by 22 university professors of physics education and 45 physics teachers. They rated each item on a 4-step Likert Scale (highly relevant, relevant, marginally relevant, and irrelevant). The ratings of these two groups were correlated with the validity-value (see step 4).

Results

Quality of the coding process

For the second step, six coders show a satisfying agreement in deciding whether a concept is scientific or non-scientific (N=200; Kappa: $.712<\varkappa<.910$; $p<.001$). The results for the third step, the mapping agreement, are reported in percentage of agreement. From the same text, all six coders identify 90% of the concepts in agreement and coded 90% of the propositions correctly (approx. 220 concepts and 590 propositions). A proposition indicates a connection between two concepts. Because some propositions have the same linking words or sentences these linkages should also be compared: Out of approx. 350 different linkages the coders had a mean agreement of 70%.

Thirty videotaped lessons are analyzed by counting the used concepts. This is correlated with the centrality of the concepts out of the concept maps, what means the absolute number of incoming and outgoing connections to one concept in the map. The lessons only concern a special topic in electricity and the concept maps concern the whole content in electricity that is taught from the seventh grade to the tenth grade. So the correlation ($r=.428$; $p<.001$) is satisfactorily high.

Descriptive Results

One textbook contains between 740 and 1090 concepts and between 1400 and 1700 different linkages with up to 5000 propositions on approx. 300 pages. Even if they differ in the amount of used physics concepts, the concepts with the highest centrality are almost the same in all textbooks. As already Merzyn (1994) has stated, the authors often use different examples or synonyms. But (within the domain of physics) clearly defined concepts as "force" or "current" are used by all authors. Exemplarily in one book the most central concept is "force" with 112 connections, followed by "body" and "energy. The average number of connections per concept is 5.8; the average number of concepts per proposition is 2.9. Each side of a book mentions statistically 2.2 concepts for the first time and concludes 4.3 propositions. In approximately 15 % of the propositions, only one concept is mentioned; so there is no connection to other knowledge.

Rating on the content validity of the items

Three different ratings were compared: 45 teachers as experts, 22 professors for science education as experts, and a concept map analysis. The professors were asked for curricular validity, the teachers for curricular validity and validity with respect to their individual teaching. The concept map analysis uses at first centrality as a measure of curricular validity. Medians for all items are calculated within both expert groups. While the group of teachers is consistent in their ratings (ICC(2,2) = .942; $p < .001$ for the validity with respect to their teaching and ICC(2,2) = .903; $p < .001$ for the curricular validity), the interrater agreement for the professors is very low (intra class correlation: ICC(2,2) = .466; $p < .050$). For the third rating the NES items have to be analyzed based on the concept maps. Using the centrality-value, it was possible to calculate a validity-value for each item.

Table 1: Correlations between different content validity measurements

	teachers rating on curricular validity	*professors rating on curricular validity*	*rating based on the concept maps*
teachers rating on validity with respect to their teaching	Spearman's ϱ = .660; p = .007	Spearman's ϱ = .587; p = .021	Spearman's ϱ = .575; p = .025

The teachers rating on validity with respect to their teaching is seen as the most reasonable approach to measure content validity; on the one hand it is the expert rating with the highest interrater reliability, on the other hand it has the closest connection to what is taught in schools. As the table shows, the teachers see a difference between the curriculum and the content of their teaching. There is nearly the same agreement between the concept-map-approach and the teachers in comparison to the professors and the teachers. From all correlations between the mapping and the expert ratings the reported one (see table) was the highest.

Conclusions and Implications

The mapping procedure could be shown as valid and reliable. By then we are able to describe the structure of the content that is taught in schools in an easy and measurable way. The first and most intuitive approach to the measurement of content validity, using the centrality as descriptor for the concepts, lead to sufficient results. Within this study it could be shown that the interrater agreement of the group of professors on the curricular validity of test items was bad. This is comparable to a lot of other studies and encourages using other ways to rate content validity of test items. The new rating approach based on the textbooks is at least as precise as asking a group of professors without the problem of interrater reliability.

Understanding textbooks as a source for knowing what is taught without asking anybody and analyzing them with the developed method allows concentrating on a-priori curricular valid item development. As content validity is known as central problem in large-scale assessment, this approach will lead to a new way of judging content validity of items regarding a defined area of content. The measure can be integrated into development of each test helping to judge tests from a science education point of view in addition to the classical methodologist view.

References

American Association for the Advancement of Science (Eds.) (2007). Getting Assessment Right – Rigorous Criteria, Student Data Inform Project 2061'a Assessment Design. 2061 Today – Science Literacy for a Changing Future, 17 (1), 1 & 4-5.

Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W. et al. (2001). PISA 2000. Basiskompetenzen von Schülerinnen und Schülern im internationalen Vergleich. (PISA 2000. Students' fundamental competences in an international comparison). Opladen: Leske + Budrich.

Chiappetta, E. L., Fillman, D. A. & Sethna, G. H. (1991). A Method to Quantify Major Themes of Scientific Literacy in Science Textbooks. *Journal of research in science teaching*, *28* (8), 713 – 725.

Commons, M. L., Goodheart, E. A., Pekker, A., Dawson, T. L. Draney, K., & Adams, K. M. (2007). Using Rasch Scaled Stage Scores To Validate Orders of Hierarchical Complexity of Balance Beam Task Sequences. In: E. V. Smith & R. M. Smith (Eds.), Rasch Measurement: advanced and specialized applications (p. 121-147). Maple Grove, USA: JAM press.

Haladyna, T. M. (2004). Developing and Validating Multiple-Choice Test Items. New Jersey: Lawrence Erlbaum Associate Publishers.

Heller, K. A., & Perleth, C. (2000). Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision (KFT 4-12+ R). (Test for cognitive skills from grade 4 to 12, revision). Göttingen: Beltz Test GmbH.

Kauertz, A. (2008). Schwierigkeitserzeugende Merkmale physikalischer Leistungstestaufgaben (difficulty causing characteristics of physical test items). Berlin: Logos Verlag.

Kauertz, A., & Fischer, H. E. (2006). Assessing Students' Level of Knowledge and Analysing the Reasons for Learning Difficulties in Physics by Rasch Analysis. In: X. Liu & W. J. Boone (Eds.), Applications of Rasch Measurement in Science Education (p. 212-246). Maple Grove, USA: JAM press.

Kirk, M., Mathews, C. E. & Kurtts, S. (2001). The Trouble with Textbooks. *The Science teacher*, *68* (12), 42 – 45.

KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland [The Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany]). (2005a). Bildungsstandards im Fach Chemie für den Mittleren Bildungsabschluss. (Educational standards in chemistry for the general educational schoo- leaving certificate after grade 10). München, Neuwied, Luchterhand.

KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland [The Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany]). (2005b). Bildungsstandards im Fach Physik für den Mittleren Bildungsabschluss. (Educational standards in physics for the general educational schoo- leaving certificate after grade 10). München, Neuwied, Luchterhand.

Köller, O., Fischer, H.E., Mayer, J., Sumfleth, E., Hartmann, S., Katzenbach, M. et al. (2008). Hinweise zur Erstellung von Testaufgaben für das Projekt Evaluation der Bildungsstandards in den Fächern Biologie, Chemie und Physik, Bd. 1-5 (Information about the development of test items for the evaluation of the educational standards for the subjects biology, chemistry and physics, volume 1-5). Unpublished working paper.

Labudde, P. (2008). Developing and Implementing New National Standards in Science Education: The Role of Science Educators. In: B. Ralle & I. Eilks (eds.): Promoting successful science learning – The worth of science education research (p. 63-73). Aachen: Shaker.

Lissitz, R. W. & Samuelsen, K. (2006). Dialogue on Validity – A suggested change in Terminology and Emphasis Regarding Validity and Education. *Educational Researcher*, *36* (8), 437 – 448.

McClelland, David C. (1973). Testing for Competence Rather than for "Intelligence". American Psychologist, 28, 1, 1-14.

Merzyn, G. (1994). *Physikschulbücher, Physiklehrer und Physikunterricht.* Kiel, Germany: IPN.

Neumann, K., Trendel, G. & Fischer, H. E. (2006). Vernetzungspotential des Lehrplans Naturwissenschaft in Nordrhein-Westfalen. In Pitton, A., Ed., *Lehren und Lernen mit neuen Medien. Jahrestagung der GDCP in Paderborn 2005*. Münster, Germany: Lit Verlag, 132 – 134.

OECD [ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT] (1999). Measuring student knowledge and skills. A New Framework for Assessment. Paris: OECD Publication Service.

Riffert, F. (2005). The Use and Misunderstanding of Standardized Testing: A Whiteheadian Point of View. *Interchange*, *36* (1-2), 231 – 252.

Schecker, H. & Parchmann, I. (2007). Standards and Competence Models: The German Situation. In Waddington, D., Nentwig, P. & Schanze, S., Eds., *Making it comparable – Standards in science Education.*Münster, Germany: Waxman, 147 – 164.

Shavelson, R. J. (1971). Some Aspects of the Relationship between Content Structure and Cognitive Structure in Physics Education. Ann Arbor, U.S.A.: University Microfilms.

Standing Conference of the Ministers of Education and Cultural Affairs. (2005) [KMK]. *Bildungsstandards im Fach Physik für den mittleren Schulabschluss (NES – Physics)*. München & Neuwied, Germany: Luchterhand.

Valverde, G. A., Bianchi, L.J., Wolfe, R. G., Schmidt, W. H. & Houang, R. T. (2002). According to the book. Using TIMSS do investigate the translation of policy into practice through the world of textbooks. Dordrecht: Kluwer Academic Publishers.

Vlug, K. F. M. (1997). Because Every Pupil Counts: The Success of the Pupil Monitoring System in The Netherlands. *Education and Information technologies, 2* (4), 287 – 306.

Waddington, D.; Nentwig, P.; Schanze, S. (Eds.; 2007). Making it Possible: Standards in Science Education. Münster: Waxmann.

Walpuski, M., Kampa, N., Kauertz, A., & Wellnitz, N. Evaluation der Bildungsstandards in den Naturwissenschaften. (Evaluation of the educational standards for the science subjects). MNU, (2008).

Weinert, F. E. (2001). Concept of competence – A conceptual clarification. In : Rychen, D. S. & Salganik, L. H., Eds., *Defining and Selecting Key Competencies.* Göttingen: Hogrefe and Huber, 45 – 65.

Wilson, M. (2005). Constructing Measures: An Item Response Modeling Approach. Mahwah: Lawrence Erlbaum Associates.

Wirtz, M., & Caspar, F. (2002). Beurteilerübereinstimmung und Beurteilerreliabilität. Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen (inter-rater-concord and inter-rater-reliability. Methods for analysis and improvement of the rating-reliability with categorisation system and rating scales). Göttingen: Hogrefe.

Wu, C.-C., Lee, G. C. & Lai, H.-K. (2004). Using Concept Maps to Aid Analysis of Concept Presentation in High School Computer Textbooks. *Education and Information technologies, 9* (2), 185 – 197.

Wu, M.L., Adams, R.J., Wilson, M.R., & Haldane, S.A. (2007). Acer ConQuest Version 2.0. Generalised Item Response Modeling Software. Victoria: Acer Press.

SCIENCE TEACHERS' LEVEL OF USING ALTERNATIVE ASSESSMENT AND THEIR PERCEPTIONS *

Serkan Buldur
Cumhuriyet Üniversitesi

Nilgün Tatar
Cumhuriyet Üniversitesi

Abstract

This study aimed at determining the level of use of the alternative assessment approaches by the elementary school science teachers. Observations, interviews, and document analysis were used as data collection tools in the scope of this study which adopted case study method. Study sampling was composed of 13 science teachers. The teachers who participated in the study were observed for a period of six weeks to determine how often they used alternative assessment techniques and tools; the problems they encountered when using such techniques and tools; the areas in which they succeeded and in which they failed and; the sources they used. At the end of this six-week application, teachers were interviewed to learn their opinions about the above-mentioned points. According to the analysis, performance evaluation took the first rank among the tools and techniques frequently used by the teachers and timing was the biggest problem they encountered while using these tools and techniques. While using these approaches, teachers were found to be most successful in increasing student participation in the lesson and to be most unsuccessful in using different tools and techniques. The sources most frequently used by the teachers were determined to be textbooks and workbooks. Some suggestions are made in the study on the basis of the results obtained this way.

Introduction

Science education is a national issue in many countries, and considerable educational reform resources are dedicated to improving science education, especially assessment practices (Chang & Chiu, 2005; Gott & Duggan, 2002; Liu, 2000). Assessment practices have changed radically over the years. The expansion of student-centered classrooms demands evaluation tools that appropriately assess students' learning, encourage lifelong skills, and provide teachers with insights and diagnostic information as well as enhancing teaching effectiveness (Benson & Smith, 1998).

Alternative assessment—which also is referred to as classroom-based, qualitative, informal, or performance assessment—is a way to gauge student learning other than formal testing. Alternative assessment exhibits several distinguishing characteristics:

• Alternative assessment is situated in the classroom with teachers making choices in the measures used.
• Alternative assessment is based on a constructivist view of learning whereby the student, the text, and the context impact learning outcomes.
• Alternative assessment is predicated on the view that learning processes are equal to, if not greater than, the resulting products (Janisch, Liu & Akrofi, 2007).

Alternative assessment evaluates students using a variety of methods such as portfolios, projects, self and peer assessment, checklists, performance-based outcomes, journals, concept maps, drawings, interviews, oral quizzes, Vee diagrams, rubrics, and other types of open-ended approaches. Any technique being used may assess the knowledge, skills, and attitudes being possessed by students. Effective use of such techniques and tools allows teachers to closely observe the development of their students.

Although many educators agree on the importance of using a variety of authentic assessment techniques in the classroom, implementing them is difficult. In other words, many teachers may be unsure of how to combine quality assessment with daily practice (Corcoran, Dershimer & Tichenor, 2004). In order to perform an effective use of alternative assessment approaches in classrooms, teachers ought to possess knowledge with regard to philosophical and conceptual bases of alternative assessment approaches.

Alternative assessment started to be considered under the assessment and evaluation dimension of the science and technology curriculum with the curriculum change made in Turkey in 2004. However, teachers had difficulty assessing students by alternative assessment methods. Teachers do not possess in-depth knowledge with regard to why, when, and how they make use of alternative assessment tools. Similar results have been obtained from numerous related studies.(Korkmaz and Kaptan, 2003; Al-Sadaawi, 2007; Şaşmaz-Oren and Tatar, 2008). Korkmaz and Kaptan (2003), Watt (2005) and Cheng (2006) suggested in their studies that teachers should be given trainings to ensure effective implementation of alternative evaluation. The use of alternative assessment tools requires a theoretical knowledge base about the purpose underlying their use, such as responsive teaching and children's metacognitive awareness of their own accomplishments and future learning goals (Janisch, Liu & Akrofi, 2007).

Determination of the knowledge and opinions of teachers on alternative assessment approaches will guide the researchers on their way to determine and overcome the deficiencies on this issue. New assessment approaches may thereby be introduced to teacher candidates, and teachers as well, and effective use of such assessment approaches in the classrooms may be maintained thereafter.

Rationale

The purpose of this study was to determine how often science teachers use alternative assessment and their perceptions about these assessment approaches. The main questions being asked in this study are:

During teachers use of alternative assessment;

1. Techniques and tools they have frequently used,
2. The difficulties they encountered,
3. The areas in which they succeed and in which they failed,
4. The sources they used.

Methods

Case study method was adopted in this study. Observations, interviews and document analysis were used as data collection tools in the scope of the study. Study sampling was composed of 13 science teachers employed in nine elementary schools. Table 1 show that teachers who participated this study demographic features.

In this scope, observation and interview forms were used as data collection tools. Moreover, the questions included in the written examinations made by the teachers during the semester were analyzed via document review method and were explained through descriptive analysis. Data collection tools that were used in the study were developed by the researchers. Observation form constituted of 2 dimensions. First dimension included the names of 24 alternative assessment and evaluation tools and techniques. It was developed as a four-point Likert-type scale to determine whether teachers used these techniques and tools during the observation period and, if yes, to detect how frequently they used them. The second dimension of the form included open-ended questions about the difficulties they encountered while using these techniques and tools, the areas in which they succeeded and in which they failed and the sources they used.

Standardized open-ended interview method was adopted in the study. The first part of the interview form was composed of the questions about personal information; the second part of 3-point Likert-type questions aimed at determining whether they used alternative assessment and evaluation tools and techniques and, if yes, the level of use of such tools and techniques by them; the third part of open-ended questions. The questions asked in the third part of the form aimed at assessing the situations same with those assessed via the open-ended questions in the observation form.

Implementation was carried out with 50 fourth-grade university students attending at the Department of Science Teaching and taking the lesson "School Experience II". These teacher candidates were given information on alternative assessment and evaluation techniques and tools for a six-week period. Model implementations were carried out by these teacher candidates about how to prepare and use these techniques and tools. They made observations to determine the level of use of the alternative assessment and evaluation techniques and tools by the science teachers employed in these nine primary schools. They observed the teachers for a period of six weeks on 2 lessons/week basis. At the end of the observation period, interviews were conducted with these 13 science teachers. The data obtained via observation and interview forms were analyzed by the researchers through descriptive analysis. In addition, the questions included in the written examinations made by the teachers in the semester were collected to determine whether the teachers used alternative assessment and evaluation tools and techniques in these questions.

Table 1. Teachers' demographic characteristics about various variables

Teachers	Gender	Experience	In-service training about alternative assessment
T_1	Male	10 Years	-
T_2	Male	25 Years	-
T_3	Male	13 Years	-
T_4	Female	16 Years	-
T_5	Female	15 Years	+
T_6	Male	15 Years	+
T_7	Male	28 Years	+
T_8	Male	10 Years	+
T_9	Female	10 Years	+
T_{10}	Male	15 Years	+
T_{11}	Male	10 Years	+
T_{12}	Male	24 Years	+
T_{13}	Female	5 Years	+

Results

In this chapter, data derived from observation and interviews are explained in relation with the problems of the study.

Alternative assessment techniques and tools being frequently made use by science teachers

Frequency of use of alternative assessment approaches by teachers in their classes, being observed for a period of six weeks, is displayed on Table 2. 13 science teachers were observed by 4-5 teacher candidates for 2 hours a week in a total period of six weeks. Teachers were scored by teacher candidates on whether they had used 24 techniques and tools, and if so, how often they had used in their classroom. While a technique having never been used was scored by 0 point, insufficient level was scored by 1, medium level was scored by 2, and sufficient level was scored by 3 points. Having calculated weekly scores of each teacher, and figured out their arithmetical averages, researchers, thereafter, ascertained the level of use of each technique and tool by the teachers.

Table 2. Utilization ratios of techniques and tools by teachers according to the data from observation form

	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}	T_{13}	**Mean**
Performance Evaluation	1,59	1,13	0,5	0,91	0,85	0,1	2,46	0	0,27	0	0,18	1,29	0	**0,71**
Drawings	0,57	0,89	0,23	0,58	0	0,1	1,71	0	0,58	0	0,11	1,29	0,5	**0,50**
Concept Map	0,38	0,5	1,28	0,26	0,5	0,28	0,17	0,75	0,56	0	0,05	1,5	0	**0,48**
Portfolio	0,16	0,13	0,37	1	0,75	0,1	2,21	0	0,44	0	0,06	1	0	**0,48**
Project	0,81	0,23	0,54	0,9	0,75	0,3	0,46	0	0,81	0	0,38	0,63	0	**0,45**
Poster	0,75	0,16	0,17	0,15	0,5	0,2	1,08	0	0,79	0	0,05	1,08	0	**0,38**
Word Association	0,81	0,47	0	0,33	0,1	0	1,96	0	0	0	0	0,88	0	**0,35**
Science stories & story maps	0,22	0,92	0,04	1,25	0	0,2	0	0	0,56	0	0,25	0,63	0,38	**0,34**
Self-evaluation	0,16	0,34	0,21	0,16	0	0	2,46	0	0,17	0	0	0,54	0	**0,31**
Puzzles	0	0,88	0,08	0,1	0	0,3	0	0,19	0,79	0,25	0	1,38	0	**0,30**
Semantic features analysis	1,06	1,66	0,04	0,1	0,25	0,2	0	0	0	0	0,26	0,13	0	**0,28**
Essay	0	0,06	0,29	0,43	0	0	1,5	0	0,06	0	0	0,92	0	**0,25**
Rubric	0,31	0,03	0	0,19	0,5	0	1,42	0	0,17	0	0	0,42	0	**0,23**
Interview	0,16	0,67	0,25	0,15	0,2	0,25	0	0	0	0	0	0,75	0	**0,19**
Peer Evaluation	0	0,23	0,08	0,18	0	0	1,38	0,19	0,29	0	0	0	0	**0,18**
Diagnostic tree	0	0,19	0,18	0	0,25	0,2	0	0	0	0	0	1,25	0	**0,16**
K-W-L Chart	1,09	0,13	0	0	0	0,3	0	0	0,33	0	0	0,13	0	**0,15**
Grid	0,25	0,38	0	0	0	0	0	0	0	0	0	0,67	0	**0,10**
Journal	0,06	0	0	0,05	0	0	0,25	0	0,13	0	0	0,63	0	**0,08**
Conceptual cartoon	0,13	0,06	0	0	0	0	0,17	0	0,5	0	0,2	0	0	**0,08**
Attitude scale	0	0	0,04	0	0,15	0	0,75	0	0	0	0	0	0	**0,07**
Prediction card	0,48	0	0	0	0	0	0	0	0	0	0	0,13	0	**0,04**
Flash card	0,06	0	0	0	0	0	0	0	0	0	0,2	0,25	0	**0,03**
Vee Diagram	0	0	0,04	0	0	0	0	0	0	0	0	0	0	**0,01**

According to observation data; performance evaluation, drawings, portfolio, concept maps and projects were the techniques and tools most frequently used by the teachers. On the other hand, teachers stated in the interviews that they most frequently used the techniques and tools of performance evaluation, concept maps, projects, K-W-L charts and portfolios. Examination of the observation data revealed that teachers used these techniques and tools at a low and insufficient level during the lesson. However, teachers stated in the interviews that they used these techniques and tools more frequently. This difference can be explained via the impossibility of observing all the lessons given by a teacher in a whole semester or the exaggerated statements made by teachers about the level of use of these techniques and tools. Comparisons between the observation and the interview data put forward that teachers most frequently used the techniques and tools of performance evaluation, concept maps, projects and portfolio and least frequently the techniques and tools of Vee Diagrams, attitude scales and story and comment cards. Frequent use of performance evaluation, concept maps, projects and portfolios -which are important in monitoring of the student improvement-, was regarded a positive result; however, low level of use of the other techniques and tools showed that teachers could not use all the alternative assessment and evaluation tools and techniques in an effective manner.

Difficulties being faced by science teachers, while they are making use of alternative assessment techniques and tools

Each teacher candidate pointed out the difficulties, having been encountered by the teacher he/she observed. Researchers assessed the data collected for the teachers. Difficulties, having been encountered by the teachers, are displayed on Table 3.

Table 3. Difficulties, having been encountered by the teachers, according to observation form

	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}	T_{13}	**Total**	**%**
Classroom management	11	8	0	6	5	0	2	1	4	0	1	3	0	**41**	**31,54**
Time management	8	1	4	3	0	0	0	0	2	0	1	2	0	**21**	**16,15**
Over-crowded class	2	0	4	0	4	2	3	0	0	0	0	2	0	**17**	**13,08**
Lack of teacher knowledge about alternative assessment	1	2	0	4	4	1	2	0	0	0	2	0	0	**16**	**12,31**
Lack of student interest	1	0	0	3	0	0	4	0	0	0	0	1	0	**9**	**6,92**
Lack of class facilities	1	0	1	1	0	0	3	0	0	0	0	0	0	**6**	**4,62**
Economic difficulties	4	0	0	0	0	1	0	0	0	0	0	0	0	**5**	**3,85**
Lack of tools	1	3	0	0	0	0	1	0	0	0	0	0	0	**5**	**3,85**
Lack of student level of readiness	0	1	0	0	0	1	0	0	0	0	0	0	0	**2**	**1,54**
Guiding for students	0	1	1	0	0	0	0	0	0	0	0	0	0	**2**	**1,54**
Being excited students	0	0	0	0	0	0	0	0	2	0	0	0	0	**2**	**1,54**
Communicate with students	0	0	0	0	0	0	0	0	1	0	0	0	0	**1**	**0,77**

According to observation data teachers have experienced following difficulties; classroom management, timing management, over-crowded classrooms, lack of teacher's knowledge lack of student interest, classroom order, and lack of student knowledge and skills can be listed as the difficulties encountered by the teachers while using alternative assessment and evaluation techniques and tools. All of the teachers stated during the interviews that these tools and techniques took too much time. In addition, insufficient classroom and school facilities, lack of teacher's knowledge about how to use these techniques and tools, parent complaints, the heavy content of the curriculum, over-crowded classrooms, class management, lack of students' knowledge and skills, lack of student interest and lack of sufficient sources were listed as the other difficulties encountered while using these techniques and tools.

Fields of success and failure of science teachers, while they were making use of alternative assessment techniques and tools

According to observation results, fields of success of teachers, while making use of alternative assessment tools are seen on Table 4.

Table 4. Fields of success of teachers according to observation form data

	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}	T_{13}	**Total**	**%**
Attracting student interest	7	5	1	5	1	0	1	0	0	0	0	2	0	**22**	**40,74**
Relating the subject to the daily life	0	0	0	0	5	2	0	0	0	0	0	0	0	**7**	**12,96**
Increasing student participation	3	0	0	0	0	1	0	0	2	0	0	0	0	**6**	**11,11**
Mastering of the subject	0	0	1	0	0	0	0	2	0	0	0	0	0	**3**	**5,56**
Caring individual differences	0	0	0	0	0	0	2	1	0	0	0	0	0	**3**	**5,56**
Giving feedback	0	1	0	0	0	0	0	1	0	0	0	0	0	**2**	**3,70**
Using sources	0	0	0	2	0	0	0	0	0	0	0	0	0	**2**	**3,70**
Time management	0	0	0	0	0	0	0	0	2	0	0	0	0	**2**	**3,70**
Constituting democratic classroom environment	0	0	0	0	0	0	0	0	2	0	0	0	0	**2**	**3,70**
Practicing different techniques and tools	1	0	0	0	0	0	0	0	0	0	0	0	0	**1**	**1,85**
Using appropriate techniques and tools for the lesson	1	0	0	0	0	0	0	0	0	0	0	0	0	**1**	**1,85**
To be objective	0	0	1	0	0	0	0	0	0	0	0	0	0	**1**	**1,85**
Determining level of readiness	0	0	0	0	1	0	0	0	0	0	0	0	0	**1**	**1,85**
Guiding for students	0	0	0	0	0	0	0	0	0	0	0	1	0	**1**	**1,85**

According to observation data, teachers have been found successful in following areas; in attracting student attention and increasing student participation in the lesson; relating the subject to the daily life; participating the lesson all students, ,respectively. According to the interview data, on the other hand, teachers were most successful in attracting student attention and increasing student participation in the lesson; using different techniques and tools and; making objective evaluations.

According to observation results, fields of failure of teachers, while making use of alternative assessment tools, are seen on Table 5.

Table 5. Fields of failure of teachers according to observation form data

	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}	T_{13}	**Total**	**%**
Using different tools and techniques	3	3	2	1	0	3	3	0	0	0	2	5	0	**22**	**33,33**
Classroom management	2	7	0	5	0	1	0	0	0	0	0	1	0	**16**	**24,24**
Giving feedback	1	1	1	2	0	1	0	0	0	0	0	1	0	**7**	**10,61**
Increasing student participation	0	1	0	3	0	0	0	0	1	0	0	1	0	**6**	**9,09**
Using rubric	0	1	0	0	0	0	2	0	0	1	0	0	0	**4**	**6,06**
Time management	1	0	0	0	0	0	0	0	1	0	0	0	0	**2**	**3,03**
Preparing appropriate activities for student level	1	0	0	0	0	0	1	0	0	0	0	0	0	**2**	**3,03**
Guiding for students	0	0	0	0	0	0	0	0	2	0	0	0	0	**2**	**3,03**
To profit from materials	0	0	0	0	1	0	0	0	0	0	0	0	0	**1**	**1,52**
Selecting appropriate different techniques and tools for activities	0	0	0	0	0	0	1	0	0	0	0	0	0	**1**	**1,52**
Other	1	0	0	1	0	0	0	0	1	0	0	0	0	**3**	**4,55**

According to observation data teachers have failed following areas; using different tools and techniques, classroom management; providing feedback; student guidance; use of rubric and use of time, respectively. Interview data, on the other hand, showed that teachers failed in preparing and using techniques and tools; use of time; classroom management and; providing feedback

Resources being benefited by science teachers, while they are making use of alternative assessment techniques and tools

Resources, having frequently been preferred by teachers for making use of alternative assessment tools, are seen on Table 6.

Table 6. Resources, having been used by the teachers, according to observation form

	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}	T_{12}	T_{13}	**Total**	**%**
Textbooks	14	8	10	12	2	3	8	3	4	2	4	7	0	**77**	**36,49**
Workbooks	4	11	8	9	5	5	4	0	5	0	2	5	0	**58**	**27,49**
Internet	1	0	0	0	4	0	0	0	3	0	0	0	0	**8**	**3,79**
Supplementary books	1	0	13	5	4	0	0	0	1	0	0	7	0	**31**	**14,69**
Other	1	1	0	0	0	0	0	2	0	0	2	0	0	**6**	**2,84**

According to the observation results, teachers frequently preferred using the techniques and tools presented in the textbooks, workbooks and supplementary books as well as the tools they prepared. Eight teachers were observed as using the devices, having been prepared by themselves

Regarding use of sources, interviews produced data similar to those of observations. In addition, 3 teachers stated that they also used the techniques and tools they developed themselves.

Conclusions and Implications

In this study, the ways by which elementary education science teachers make use of alternative assessment approaches, as well as their opinions toward these approaches have been determined. Comparisons between the observation and the interview data put forward that teachers most frequently used the techniques and tools of performance evaluation, concept maps, projects and portfolio and least frequently the techniques and tools of Vee Diagrams, attitude scales and story and comment cards.

Similar results are encountered in numerous studies, included in the literature. In the study they carried out on 159 teachers, Culbertson and Wenfan (2003) found out that portfolio, check lists and projects were the tools frequently used by the teachers. Duban and Küçükyılmaz (2007), Güven and Eskitürk (2007) and Özdaş et al. (2007) suggested in their studies that performance evaluation was the tool most frequently used by the teachers. The reason for frequent use of performance assessment by teachers may be explained as follows. The study conducted by Tatar et al. (2008) revealed that performance evaluation was given the largest space in the textbooks and workbooks. Another result of the present study was that textbooks and workbooks took the first rank among the sources teachers used. Depending on this result, it can be suggested that sources are effective on the frequency of the use of such techniques and tools.

Analysis of the problems revealed by the observation and interview data pointed out timing as the most frequently encountered problem. The studies carried out by Thompson et all. (2001), Mintah (2003), Watt (2005) and Çalık (2007) produced similar results as well. It is known that while the preparation of these student-oriented techniques and tools take more time than the preparation of conventional assessment and evaluation techniques and tools, the former reflects student improvement better than latter. Classroom management may possibly be more difficult in the interactive lessons increasing student participation. However, in-classroom rules to be set and the determination of the teacher in applying these rules can facilitate this process.

In the light of the findings both observation form and interview form teachers most successful about attracting student’s interest in the lesson. Morgil and Aydın, (2005), Maslovaty and Kuzi (2002), Cohen (1995) and Ogan-Bekiroğlu (2008) suggested in the studies they carried out that alternative assessment techniques and tools motivated students and increased their interest in the lesson. Successful use of these techniques and tools by teachers can increase the effectiveness of such techniques and tools. The failure of teachers in using alternative assessment and evaluation techniques and tools can be explained via their lack of knowledge on this issue. Ford and Ohlhausen (1991), Korkmaz and Kaptan (2003), Watt (2005) and Cheng (2006) suggested in the studies they conducted that teachers had quite limited knowledge on these techniques and tools.

In the light of these findings, it can be concluded that there are limited number of sources on this issue. Increasing the number of the sources explaining how to prepare and use alternative assessment and evaluation techniques and tools will be beneficial in filling the knowledge gap of teachers on this issue. Unfamiliar assessment procedures and terminology can cause a lack of confidence for teachers; therefore, teachers need to be knowledgeable and comfortable with alternative assessment techniques and tools before implementing the process. For this reason, more emphasis should be given to planning and implementing alternative assessments in pre-service and in-service teacher education.

In order to implement alternative assessment approaches teachers must use various sources. Therefore additional sources must be made available for teachers in the classroom such as;, internet connection, computer, clever board etc. Opinions of teachers and/or teacher candidates towards alternative assessment may be reviewed by taking different variables (motivation, self-sufficiency, attitude, etc.) into consideration, and by making use of qualitative and quantitative research methods as well.

References

Al-Sadaawi, A. S. (2007). An Investigation of Performance-Based Assessment in Science in Saudi Primary Schools. Unpublished PhD thesis. Victoria University Melbourne-Australia.

Benson, R. T & Smith, L. J. (1998). Portfolios in First Grade: Four Teachers Learn to Use Alternative Assessment. *Early Childhood Education Journal*, Vol. 25, No. 3.

Chang, S.-N. & Chiu, M.-H. (2005). The development of authentic assessments to investigate ninth graders' scientific literacy: in the case of scientific cognition concerning the concepts of Chemistry and Physics. *International Journal of Science & Math Education*, 3(1), 117-140.

Cheng, M. H. (2006). Junior Secondary Science Teachers' Understanding and Practice of Alternative Assessment in Hong Kong: Implications for Teacher Professional Development. *Canadian Journal of Science, Mathematics and Technology Education*, 6 (3): 227-243.

Cohen, K.M. (1995). Achieving Positive Attitudes toward Science through Alternative Assessments. Unpublished Master Thesis. Master of Arts in Teaching. Graduate Faculty of Christopher Newport University.

Corcoroan, C. A., Dershimer, E. L. &Tichenor, M. S. (2004). A Teacher's Guide to Alternative Assessment Taking the First Steps. *The Clearing House*. Vol. 77, No. 5, 213-216.

Culbertson, L. D. & Wenfan, Y. (2003). Alternative Assessment: Primary Grade Literacy Teachers' Attitudes and Practices. Paper Presented at the Annual Meeting of the American Educational Research Association. (Chicago, IL, April21-15, 2003).

Duban N. & Küçükyılmaz. (2008) A. Primary Education Pre-service teacher's opinions regarding to use of alternative measurement-evaluation methods and techniques in practice school. *Elementary Education Online*, 7(3), 769-784

Gott, R. & Duggan, S. (2002). Problems with the Assessment of Performance in Practical Science: Which Way Now? *Cambridge Journal of Education*, 32(2).

Ford, M. P. & Ohlhausen N. M. (1991). Portfolio Assessment in Teacher Education Courses: Impact of students' Beliefs, Attitudes and Habits. Paper Presented at the Annual Meeting of the National Reading Conference. (Palm Springs, CA, December 3-7, 1991).

Janisch, C., Liu, X. & Akrofi, A. (2007). Implementing Alternative Assessment: Opportunities and Obstacles The Educational Forum. Vol. 71,221-230.

Maslovaty N. & Kuzi E. (2002). Promoting motivational goals through alternative or traditional assessment. *Studies in Educational Evaluation*, 28, 199–222.

Mintah, J. K. (2003). Authentic Assessment in Physical Education: Prevalence of Use And perceived Impact on Students' Self-Concept, Motivation and Skill Achievement. Measurement in Physical Education and Exercise Science, 7 (3), 161-174.

Ogan-Bekiroğlu, F. (2008). Assessing Assessment: Examination of pre-service physics teachers' attitudes towards assessment and factors affecting their attitudes. *International Journal of Science Education*, 1–39.

Shepard, L. (1995). Using assessment to improve learning. *Educational Leadership*, 52(5), 38-43.

Şaşmaz-Ören, F.& Tatar, N. (2007). İlköğretim Sınıf Öğretmenlerinin Alternatif Değerlendirme Yaklaşımlarına İlişkin Görüşleri-I (Perceptions of Classroom Teachers about Alternative Assessment). *Buca Eğitim Fakültesi Dergisi*. 22: 15-28

Thompson, S. J., Benson, S. N. K., Pachnowski L. M. & Salzman, J. A. (2001). Decision-Making in Planning and Teaching. Addision-Wesley Educational Publishers Inc., United States.

Watt, H. M. G. (2005). Attitudes to the Use of Alternative Assessment Methods in Mathematics: A Study with Secondary Mathematics Teachers in Sydney, Australia. *Educational Studies in Mathematics*, 58: 21–44.

INVESTIGATING FACTORS AFFECTING HIGH SCHOOL STUDENTS' PHYSICS SELF-EFFICACY

Haki Peşman
Fırat Üniversitesi

Sevda Yerdelen
Yüzüncü Yıl Üniversitesi

Abstract

This study investigated the effects of school type, grade level and gender on high school students' physics self-efficacy. The participants of the study were 338 students from a public high school (PHS), an Anatolian high school (AHS), and an Anatolian teacher training high school (ATTHS). The schools in this study differ from each other in terms of student admission requirements and the courses they offer. The data were analyzed using two- way ANOVA and independent t-test. The results indicated that there was a significant interaction between grade level and school type in their influences on students' physics self-efficacy. The interaction effect had also practical importance. Because of the observed significant interaction, one-way ANOVAs, the follow- up tests, were conducted for each school and each grade level separately. The findings suggested that, in PHS, students' physics self efficacy scores became higher as the grade level goes upper. However, the scores in AHS and ATTHS remain stable as the grade level changes. Moreover, the school type had an impact on students' physics self-efficacy scores in 9th and 10th grades. Finally, a significant gender difference in students' physics self-efficacy did not found in the study.

Introduction

Self-efficacy is defined as self-appraisal of one's own ability to master a task. It involves both judgment about one's ability to achieve a task and one's confidence in one's skills to perform that task (Pintrich, Smith, Garcia, and McKeachie, 1991). Bandura (as cited in Schoon & Boone, 1998) claimed that people's self- efficacy beliefs had an effect on their behaviors. Behaviors happen when they believe in their own ability to perform that behavior and when they expect considering their own life experiences. Many studies in education noted that there was positive relationship between students' self-efficacy beliefs and their learning (Pajares & Graham, 1999; Schoon & Boone, 1998; Wadsworth, Husman, Duggan, & Pennington, 2007).

Rationale

The studies investigating gender differences in students' physics self-efficacy have provided inconsistent results. For example, Cavallo, Rozman and Potter (2004) observed that male students had significantly higher ***self-efficacy*** than female students. However, Shaw (2004) found there was no gender difference in physics self-efficacy for college and university physics courses. Moreover, there is little known about the effect of school type and grade level on students' physics self-efficacy.

In this study, three different types of schools were surveyed. These schools were different from each other with respect to their curriculum and students' level of achievement. Determination of factors influencing students' self-efficacy is essential for promoting their self-efficacy because of its reported positive relationship to learning. In line with this necessity, the current study aims to investigate the effects of school type, grade level and gender on high school students' physics self-efficacy.

The research questions, investigated in the scope of the study, are as follows:

1. Is there a significant difference among physics self-efficacy scores of students in different high schools?
2. Is there a significant difference among physics self-efficacy scores of students at different grade levels?
3. Is there a significant difference in the effect of grade level on high school students' physics self-efficacy for different type schools?
4. Is there a significant difference in the effect of school type on high school students' physics self-efficacy for different grade levels?
5. Is there a significant difference between male and female students' self-efficacy scores?

Methods

This study is a cross sectional survey study, which is a quantitative research methodology. The sample included 338 students (175 females, 163 males) from three different high schools; a public high school (PHS), an Anatolian teacher training high school (ATTHS), and an Anatolian high school (AHS). PHSs accept any student graduating from junior high school while ATTHSs and AHSs accept students who are able to be successful in the national entrance examination. Thus, ATTHSs and AHSs generally admit more successful students than PHSs. ATTHSs and AHSs differ in that pedagogical courses are given to students of the ATTHSs in addition to the courses students of AHSs take. The primary purpose of ATTHs is to prepare students for the educational faculties of the universities.

The instrument of the study was the adapted Turkish version of Motivated Strategies for Learning Questionnaire (MSLQ) translated and adapted into Turkish by Sungur (2004). MSLQ developed by Pintrich, Smith, Garcia, and McKeachie (1991) includes two section as motivation and learning strategies section. These sections also contain sub-scales. Self- efficacy scale contains eight seven-point Likert-scale items. Pintrich et al. (1991) found reliability coefficient (alpha) as .93 for self- efficacy scale. Sungur (2004) also estimated alpha coefficient as .89. Because the present study aimed to investigate students' self efficacy beliefs about physics, the scale translated by Sungur was changed by replacing the subject name "biology" with "physics". Alpha coefficient was calculated in this study as .93.

Results

A two- way ANOVA was conducted to examine the impact of school type and grade level on high school students' physics self-efficacy. There were statistically significant main effect for grade level [F (2, 328)=5.08, p=.007] and the interaction effect [F (4, 328)=5.25, p=.000]. The effect size (eta squared=.030) for grade level was small. Effect size (eta squared= .060) for interaction was medium. The main effect for school type [F (2, 328)=.031, p=.97] did not reach statistical significant. It also did not have practical importance since effect size (eta squared=.00) was very small. The interaction between grade level and school type can be seen clearly from Figure 1. The change of students' physics self-efficacy scores across grade levels was not same for each school and the change of students' physics self-efficacy scores across the schools was not same for each grade level. Because the significant interaction effect was obtained, the follow-up tests were conducted to examine whether differences in self-efficacy scores among grade levels and among the schools were statistically significant. For this purpose, one- way ANOVA was carried out for each school and each grade level separately. As a result, there were no any significant differences among self-efficacy scores of students at different grade levels AHS and ATTHS while there were significant differences among self-efficacy scores of students at different grade levels in the PHS [F(2, 125)=16.47, p=.000]. Also, the effect size was very large for the PHS (eta squared=.21). In addition, when the students were divided according to grade level and the ANOVA was conducted, PHS's ninth grade students differed significantly from those of AHS and ATTHS [F(2, 117)=4.576, p=.012]. There was a medium effect size (eta-squared=.07). For 10th grades, differences were also significant [F (2, 107)=3.082, p=.05]. Effect size was medium (eta-squared=0.05). For 11th grades, differences were not statistically significant but practically significant (eta squared=.04, respectively).

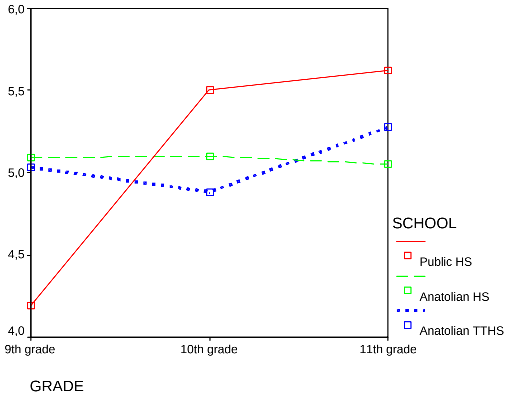


Figure 1. The plot of self-efficacy scores for PHS, AHS, and ATTHS, across three grade levels.

Finally, the independent- samples t-test was conducted to compare the mean of physics self-efficacy scores for females and males. There was no significant difference in scores for females (M=4.98, SD=1.41), and males (M=5.17, SD= 1.24); [t (335)=-1.32, p=.19]. The effect size (eta squared=.005) was very small.

Conclusions and Implications

The results showed that interaction between school type and grade level had both statistical and practical importance. In PHS, students' self efficacy scores increased across grade levels and the ninth grade students' physics self-efficacy score was significantly different from other two grade levels. On the other hand, grade level did not influence the students' self efficacy in AHS and ATTHS. The results also indicated that in the ninth grade, only PHS students and ATTHS students differed significantly from each other and the mean score for PHS students was smaller than both ATTHS and AHS students'. At 10th grade, the significant difference in physics self-efficacy scores was only between the PHS and the ATTHS. The physics self-efficacy scores of the schools did not differ for the 11th grade students. Briefly, the findings of the study showed that school type had a significant influence on change in students' physics self-efficacy across grade levels. Students' self-efficacy scores did not change across grade levels in the schools accepting more successful students while students' physics self-efficacy increased across grade levels in PHS which does not take any success in entrance examination into account. Although in grade nine which is first year in the school, students of PHS had the least physics self-efficacy scores compared to students of the other schools, in grade 11, they have the highest self-efficacy scores.

Reasons for all these differences arouse new questions. For example, "why did ninth graders of the PHS have the lowest scores?" or "why did 11th graders have the highest?" The reason for PHS's ninth grade students' lowest self-efficacy scores may be due to their failure in the national entrance examination. Answering the latter is a little bit difficult. Maybe, the differences are due to differences such as curricula or students' characteristics. However, all these possible predictions at this point are only hypotheses. That is, this study leads to future investigations of the schools in detail to identify what happen in the schools that cause this situation. The present study was a cross-sectional survey study, thus, each grade had different students. Therefore, this study also encourages the researcher to conduct longitudinal survey study with the same students to collect strong evidences about the influence of grade level and school type on students' physics self-efficacy. This study also contributes to the literature providing evidence that there is no significant difference between males' and females' physics self-efficacy.

References

Cavallo, A. M. L.Potter, W. H., & Rozman, M. (2004). Gender differences in learning constructs, shifts in learning constructs, and their relationship to course achievement in a structured inquiry, yearlong college physics course for life science majors. School Science and Mathematics, 104 (6), 288-300.

Pintrich, P. R., Smith, D. A., Garcia, T., & McKeachie, W. J. (1991). A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ).(Report No. NCRIPTAL- 91-B-004). Ann Arbor, MI: National Center for Research to Improve Postsecondary Teaching and Learning. (ERIC Document Reproduction Service No. ED 338122).

Schoon, K. J., & Boone, W. J. (1998). Self-efficacy, alternative conceptions of science, and preservice elementary teachers. Science Education, 82 (5), 553-568.

Sungur, S.(2004). An implementation of problem based learning in high school biology courses. Unpublished doctoral dissertation, Middle East Technical University, Ankara, Turkey.

Shaw, K. A. (2004). The development of a physics self-efficacy instrument for use in the introductory classroom. In J. Marx, S. Franklin, & K. Cummings (Eds.), 2003 Physics Education Research Conference: AIP Conference Proceedings, 720,173-176. Melville, NY: AIP

Pajares, F., &Graham, L. (1999). Self-efficacy, motivation constructs, and mathematics performance of entering middle school students. Contemporary Educational Psychology, 24, 124-139.

Wadsworth, L., Husman, J., Duggan, M.A., & Pennington, M. N.(2007). Online Mathematics Achievement: Effects of Learning Strategies and Self-Efficacy. Journal of Developmental Education, 30 (3), 6–14.

QUESTION POSING AND GRAPHING SKILLS OF YOUNG GIFTED STUDENTS: GENDER AND EQUAL OPPORTUNITIES ASPECTS

Dana Fischer-Shachor
Technion, Israel Institute of Technology

Miriam Carmi
Technion, Israel Institute of Technology

Yehudit Judy Dori
Technion, Israel Institute of Technology,
Massachusetts Institute of Technology, Cambridge

Abstract

Examining the status of women in developed countries, researchers indicate that at the beginning of the third *Millennium women are under-represented in high positions in many domains, including academia, industry, and the military. The definition of Giftedness is controversial and researchers differentiate between the traditional definitions Assessing higher order thinking skills can serve for identifying diverse capabilities of gifted students to comply with the multidimensional definition of giftedness. The research population included 487 gifted students, 275 of them were boys and 212 girls. The girls were further divided into those who entered the gifted program without affirmative action (N=128) and with affirmative-action (N=80). About half of the students were in the first year and the rest in their third year of participating in the gifted program. To assess students' question posing and graphing skills, we developed performance questionnaires, which included case studies. The analysis of the students' responses revealed that the gifted girls who entered the programs with affirmative action performed the same as the others or at a higher level. The results demonstrate that effective assessment can be based on performance questionnaires that had been validated.*

Introduction

The definition of *Giftedness* is controversial and researchers differentiate between the traditional definition, which suggests that giftedness equals high intelligence as measured by I.Q. (Terman & Oden, 1959) and the modern, multidimensional definition, which characterizes intelligence via cognitive, qualitative, psychological, and social aspects (Sternberg, 1984; Nevo & Chawarski, 1997; Renzulli, 1978; Tannenbaum, 1983; Gardner, 1982; 1983). In Israel, identifying gifted students follows the traditional definition and is based on I.Q. tests (Silverman, 1986). Assessing higher order thinking skills (Dori, 2003; 2007, Zohar & Dori, 2003) can serve for identifying diverse capabilities of gifted students to comply with the multidimensional definition of giftedness. This research is aimed at examining what differences, if any, exist between sub-groups[1] of gifted students in question posing and graphing skills, as reflected in their responses to case-based questionnaires?

Rationale

Examining the status of women indicates that at the beginning of the third Millennium women are under-represented in high positions in many domains, including academia, industry, and the military (AIP, NSF). Similarly, gifted girls are also under-represented in all the educational frameworks for advancing gifted children (Zormam & David, 2000). Committed to equity-based education and striving to improve this situation, the Israeli Ministry of Education has carried out a policy of affirmative action through the acceptance process for gifted girls, in some of the pull out programs and enrichment classes.

Methods

The research population included 487 Israeli gifted students, 275 of them were boys and 212 girls, with girls further divided into those who entered the gifted program without affirmative action (N=128) and with affirmative action (N=80[2]). About half of the students were in the first year and the rest in their third year of the gifted program. To assess students' question posing and graphing skills, we developed performance questionnaires, which included case studies (short stories based on students' everyday life) along with assignments which examine these skills. One version of the questionnaire was developed for young 3rd and 4th graders (first year in the program) and another version—for 5th and 6th graders (third year in the program). Initially, we validated the questionnaires by administering them to gifted students who studied in similar gifted programs and did not participate in the main study. Analysis of the responses also served for improving these questionnaires. 'Snack Time' is an example of a case-based paragraph followed by an assignment designed to examine students' question posing skills.

Snack Time

Many children like to have snacks, as deserts, between meals, on a field trip, or during recess. The snack options are unlimited – they can be sweet or salty, cubic or cylindrical, large or small, in personal or family-size bags, single- or multi-

Having you read the **Snack Time** story, what made you curious? What else would you like to know about the subject discussed here? Please write at least two questions.

Table 1 presents a sample of questions students posed and their scoring based on several categories.

The rubric included the next categories:
Complexity –Refers to how much elaborated is the question included in the response. A low level is when the response to the question is included in the story, a medium level is the level of knowledge and a high level of complexity includes several domains of interest.
Phrasing- Refers to the way the response was expressed, and if it was vague and non clear or well written.

---

[1]Boys, girls who were accepted without affirmative action, and girls who were accepted to the gifted programs with affirmative action
[2] We have no exact information about 4 girls

Table 1. Examples of two students' responses to the posing questions assignment

Grade	Quest. #	Questions posed by the students (their responses)	Complexity	Phrasing	Relation between questions	Total score
6th grader	1	Which snack is the winning snack?	3	2	3	21
	2	In which state do kids eat the largest amount of snacks?	3	2		
	3	What is the newest snack?	3	1		
	4	What is the oldest snack?	3	1		
3rd grader	1	Are there any snacks that make you gain weight or less?	2	1	3	10
	2	Are there any bitter snacks?	2	2		

An assignment designed to examine graphing skills dealt the next story:

Students from the 3rd grade were asked by their teacher in which after-school sports activities they participated. She wrote on the board in class the answers. The results were: swimming – 3 boys and 5 girls, basketball – 4 boys and 4 girls, and tennis – 6

The assignment that followed was:
"Draw a graph which describes the number of boys and girls combined who participated in the after school activities in your class."
Figure 1 and 2 present the responses of Student A and B and Table 2 is the corresponding rubric

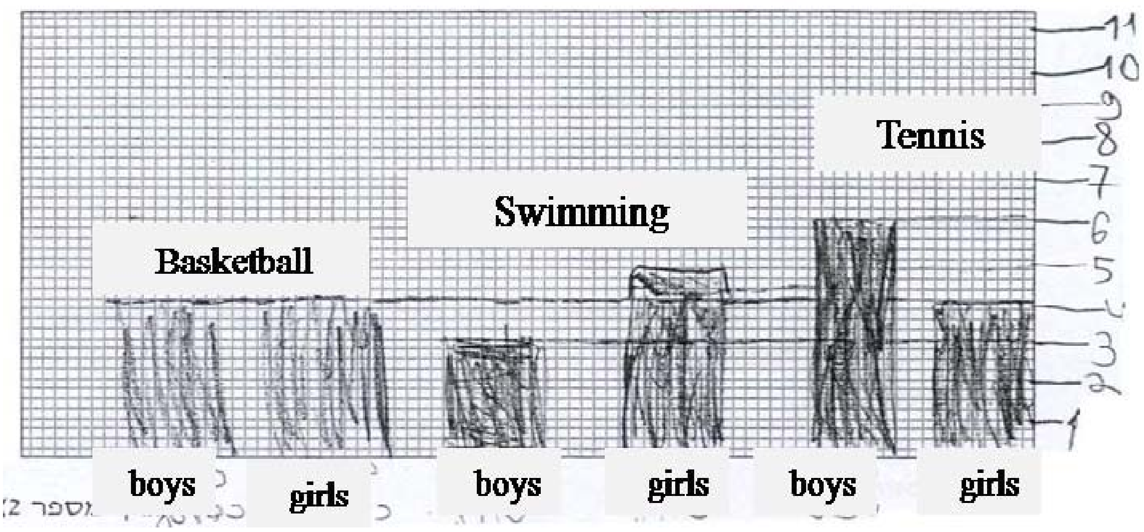


Figure 1. Student A response to the graphing skills question

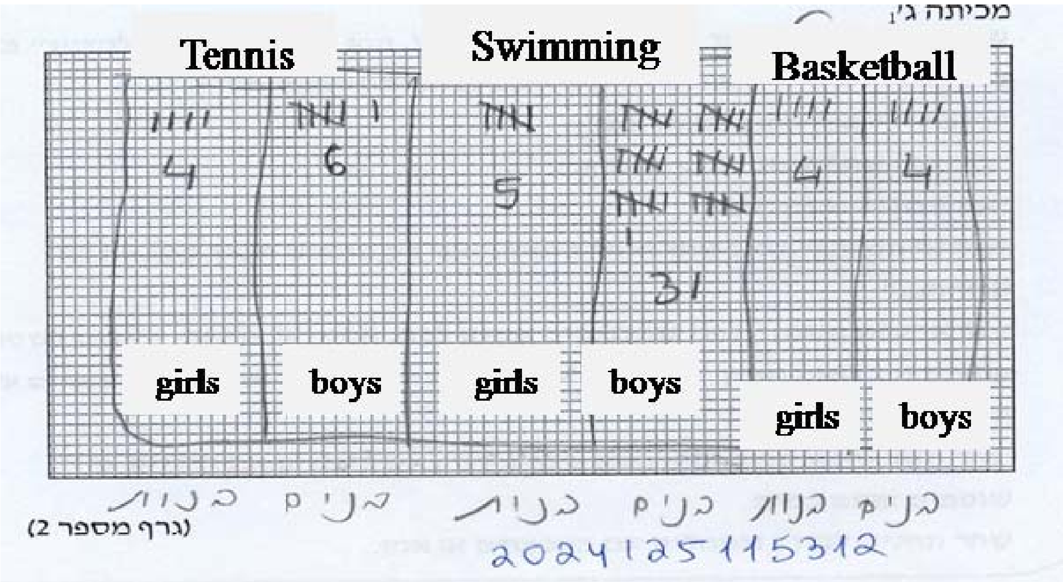


Figure 2. Student B response to the graphing skills question

Table 2. Rubric for Student A response to the graphing skills question

Total (7/7)	Did the student include scale in his graph? (1)	Did the student draw appropriate graph? (2)	Did the student process the data? (2)	Did the student answer this question? (2)	student
5/7	1	1	1	2	A
2/7	0	0	1	1	B

Results

Analysis of the responses served as a basis for improving and validating the performance questionnaires. The analysis of pre& post test responses of the young students sorted by subgroups to all the graphing skills questions are presented in Table 3.

Table 3. Analysis of the pre& post test responses of the young students - graphing skills

stage	Research sub- groups	N	$\bar{X}$	S.E.	d.f	F	<p
pre	Boys	150	70.5	1.47	258	6.8	<0.001
	Girls without affirmative action	65	61.2	2.56			
	Girls with affirmative action	46	72.4	2.70			
	Total	261	68.5	1.18			
post	Boys	110	69.4	1.9	195	0.52	n.s
	Girls without affirmative action	49	67.8	2.4			
	Girls with affirmative action	39	72.0	3.2			
	Total	198	69.5	1.4			

There were no significant differences between the scores of Girls who were accepted to the gifted programs with affirmative action and the boys. The scores of the Girls who were accepted to the gifted programs without affirmative action were significant lower than the scores of the Girls who were accepted with affirmative action and the boys.

Analyzing the posttest results of the same students after a year (see table 3) revealed no significant difference in the mean scores between these subgroups.

Similar results were obtained for the answers of the grown up students who studied two years in the gifted program, i.e., the significant differences we had found in graphing skills between students' subgroups in the beginning of the gifted program were no longer detected after two years of participation. The performance of grown-up students was significantly higher than that of the young ones, but no significant differences between the subgroups was found

Analyzing the assignments aimed at examining posing question and asking inquiry questions skills, we found no significant difference in the pretest mean scores between the young student's subgroups. However, the results of the same questions for the students' responds, as presented in Table 4, and the young students' posttest (after they completed their 2[ed] year in the gifted program) show a significant difference in the mean scores between the sub-groups. As Table 4 shows, the scores of the girls who were accepted to the gifted program without affirmative action were significantly higher than those of the boys. There was no significant difference in the mean scores between the girls who were accepted to the gifted program without affirmative action and those who were accepted with affirmative action.

Table 4. Analysis of the grown up students pretest responses to assignments related to question posing and asking inquiry questions

Stage	Research sub- groups		N	$\bar{X}$	S.E.	d.f	F	<p
Pretest	young	Boys	150	56.4	1.7	258	0.56	n.s
		Girls without affirmative action	65	59.8	2.5			
		Girls with affirmative action	46	57.1	3.2			
		Total	261	57.4	1.3			
Posttest	young	Boys	110	64.1	2.1	195	6.43	<0.0001
		Girls without affirmative action	49	76.8	2.3			
		Girls with affirmative action	39	68.0	3.4			
		Total	198	68.0	1.5			
	grown up	Boys	125	70.2	1.7	219	3.1	<0.05
		Girls without affirmative action	63	77.3	2.2			
		Girls with affirmative action	34	72.0	3.0			
		Total	222	72.5	1.2			

Conclusions and Implications

Findings indicate that the performance of the grown up students in both examined skills was significantly higher than that of the young ones. The interesting outcome is that the scores of the girls who were accepted to the gifted program with affirmative action were significantly higher than those of the girls who were accepted to the gifted program without affirmative action. Our findings clearly indicate that achievements with respect to question posing and graphing skills of girls who were accepted to the gifted program due to the policy of affirmative action are not lower than the achievements of other subgroups of students in the gifted program.

In most of the thinking skills assignments the gifted girls outperformed the gifted boys. The research did not find the sub-group of girls who got to the gifted programs with affirmative action as different from the other sub-groups, their performance was in the same or even higher level than the others. The research demonstrates that effective assessment can be based on performance questionnaires that had been validated. When properly applied, these assessment tools successfully identify differences in gender and program characteristics of gifted students' higher order thinking skills, such as include question posing and graphing skills.

Acknowledgement

This study was partially funded by the Chief Scientist of the Israeli Ministry of Education and by the Jacobs Fellowship for Graduate Students at the Technion.
The opinions expressed in this paper reflect the views of the authors only.

References

Dori, Y.J. (2007).educational reform at MIT: Advancing and evaluating technologcly-base projects on- and of-campus Journal of Science Education and Technology. 16(4)

Dori, Y.J. (2003). From nationwide standardized testing to school-based alternative embedded assessment in Israel: Students' performance in the "Matriculation 2000" Project. Journal of Research in Science Teaching, 40(1), 34-52.

AIP http://www.aip.org/statistics/trends/gendertrends.html

Gardner, H. (1982). Art, mind, and brain: a cognitive approach to creativity New York: Basic Books.

Gardner, H. (1983) Frames of mind: the theory of multiple intelligences. New York: Basic Books.

Nevo, B. & Chawarski, M.C. (1997). Individual differences in practical intelligence. Intelligence, 25, 83-92.

NSF http://www.nsf.gov/statistics/

Renzulli, J.S. (1978). What makes giftedness? Reexamining a definition. Phi Delta Kappa, 60, 18-24.

Silverman, L. K. (1986). The IQ controversy- conceptions and misconceptions. Roeper Review, 8, 136-140.

Sternberg, R. J. (1984). What should intelligence tests test? Implications of a triarchic theory of intelligence for intelligence testing. Educational Researcher, 13, 5-15.

Tannenbaum, A (1983) Gifted Children: Psychological and Educational Perspectives, New York: Macmillan Publishing Co., Inc.

Terman, L.M. & Oden, M.H. (1959) The Gifted Group at Mid-Life. Stanford, Ca: Stanford University Press.

Zohar, A. & Dori, Y.J. (2003). Higher order thinking skills and low achieving students: Are they mutually exclusive? The Journal of the Learning Sciences, 12, 145-182.

Zorman R & David H. (2000) There is another way girls and women: achievements and challenges, Jerusalem "Achva" Publications (in Hebrew).

THE GRADUAL APPROACH OF THE NATURE AND ROLE OF MODELS AS MEANS TO ENHANCE 5TH GRADE STUDENTS' EPISTEMOLOGICAL AWARENESS

Anastasios Zoupidis
University of Western Macedonia

Dimitris Pnevmatikos
University of Western Macedonia

Anna Spyrtou
University of Western Macedonia

Petros Kariotoglou
University of Western Macedonia

Abstract

In this paper we present results from a research effort that supported the design and implementation of a teaching/learning sequence (TLS) on density as a property of materials to explain floating and sinking (f/s) phenomena. One of the theoretical bases for the design of the TLS was focused on model-based learning approach. The TLS has been designed and developed by researchers working in close collaboration with primary school teachers comprising a University – school partnership. Its implementation has taken place in two science classrooms of forty one (41) fifth graders (10 –11 year old students) in total. We present the basic design principles of the TLS, concerning models, as well as its theoretical roots. Moreover, we present results on students' ability to recognize aspects of the nature and role of models, which was negotiated as part of the TLS through a set of activities with an increasing level of models' abstraction. According to the results there was a shift of the students to more scientific knowledge. On the other hand, it seems that aspects of epistemological awareness (the explanatory and predictive role) are difficult for students of that age range, obstructing them from an integrated view of scientific practice.

Introduction

Models (namely a representation of an object, a concept, a process or a phenomenon, Halloun, 2004) are an important part of any scientific processes (Nersessian, 2008) and one of the most significant aspects of scientific epistemological knowledge (Treagust, Chittleborough, Mamiala, 2002). Models could be examined in at least three levels: the external models, the instructional models, as well as the mental models that are constructed in the minds both of scientists and students. Models, at all levels, are analog representations of the reality and not its copy, they serve as a tool and not as exemplar and finally, their main role is to explain and predict (Treagust et al., 2002).

Models considered as one of the facilitators of conceptual understanding and achievement in school settings, because of their importance for the development of metaconceptual awareness, metacognitive skills and intentional learning (Vosniadou, 2007). Despite the usefulness of the models, there is some evidence for the difficulties in the instruction that could be arise from the use of models. Treagust et al. (2002), for example, argued that the way students consider models, as a precise representation or an imprecise representation helps to explain some of the conflicting ideas that students may have about scientific models. For example, students that consider models as a precise and not as an analog imprecise representation (i.e. as a replica) are constrained to understand the concept of scientific model. Besides, it is known that students in primary school mainly hold a recreational view for the models (Gilbert, 1991). That is, not all students interpret the term 'scientific model' in the same way and their interpretation depends on their experiences and personal understandings. Thus, it is important to examine how students do consider models when they experience a model presentation during the instruction and find a way to avoid misconceptions arising from the models are used. Students' experience with general models seems to be the starting point in their understanding of scientific models. General models more commonly fit into the category of scale replica, whereas scientific models assume many forms and are used more analytically. By highlighting these subtle differences between different types of models, they may be used more effectively in teaching and learning science (Treagust et al. 2002). In other words, the nature and the role of models are important issues to be discussed and explicitly taught in the classroom.

For the most effective use of models in instruction, Gentner (1983) proposed a continuum of instructional models; physical microcosms (i.e. solar system), representational systems (i.e. maps, diagrams), syntactic models (i.e. analogies) and hypothetical – deductive models (i.e. gas model). Petrosino (2003) argued that it will be most fruitful to introduce students to modeling practices through models that preserve resemblance, because these models more readily sustain mappings between the model and the world. So, as students learn over a number of cases that resemblance is less fundamental than function, they become increasingly prepared to work with models that do not preserve similarity between the model and the modeled world.

Rationale

More research is needed to understand how the use of models in instructional settings could be effective and able to facilitate conceptual understanding. Vosniadou (2007) argued that "conceptual change cannot be achieved without the development of intentional learners who have the metaconceptual awareness needed to understand the difference between their naïve beliefs and the scientific concepts to which they are exposed, and are capable of producing the top-down, deliberate and intentional mechanisms that scientists use for hypothesis testing and conscious belief-revision". Thus, taking into the account that models are an important part of any scientific processes (Nersessian, 2008) and one of the most significant aspects of scientific epistemological knowledge (Treagust et al., 2002) we claim that the explicit teaching of the role of models should facilitate conceptual understanding. In that respect, the research question was: To what extent do students acquire basic aspects of the nature and role of models and abandon the recreational views? We hypothesized that, primary school children before the intervention would hold a recreational view for the models, while after the implementation their ideas would shift to more scientific ones.

Methods

Participants

This study has been conducted in two fifth-grade science classrooms in the 1st Primary school of Florina, Greece, during April 2008. Forty one (26 were boys) fifth graders (mean age 10.7-year-old), coming from two classes have participated in the study. The implementation conducted by the usual science teacher of the two classes. She is very experienced in teaching science, and cooperated with the research group in producing the implementation.

Implementation's Design

Because it is difficult, if not impossible, to teach epistemological awareness out of a specific content, the concept of density and the floating/sinking (f/s) phenomena were used as a framework to introduce the epistemological awareness about models (Treagust et al., 2002) and Control of Variables Strategy (CVS) (Boudreaux, Shaffer, Heron, & McDermott, 2008) respectively. This development followed a cyclical evolutionary process of design, implementation, analysis, validation, and redesign. Five units, each lasted for 2 teaching periods of approximately 80 min in total, were developed. This design has been based on other relevant cases in the science education literature such as didactical structures (Linjse, 1995), educational reconstruction (Duit, 1999) and Teaching Learning Sequences (TLSs) (Kariotoglou, Psillos, & Tselfes, 2001; Méheut & Psillos, 2004). These approaches are based mostly on content transformation, students' conceptions and constructivist perspectives of teaching and learning. Lately, Integrated Constructivism (i.e. constructivism giving emphasis both to pedagogical and epistemic aspects, scaffolded by aspects of interest and motivation of learners, social significance of knowledge and nature of science) emerges a new approach of constructivism in the way TLSs are implemented (Meheut and Psillos 2004). Taking into account these approaches a TLS is developed for the study of both declarative (density of materials found easily in the environment like a piece of wood, glass, plastic, etc., as well as hi-tech materials like fiber optics or carbon) and procedural knowledge. More specifically, the procedural knowledge that students were taught were both, a) CVS method, and b) modeling skills as well as aspects of the nature and the role of models.

Furthermore, the difficulties usually children face in understanding floating/sinking phenomena and their misconceptions for the concept of density (Fassoulopoulos, Kariotoglou, & Koumaras, 2003; Havu, 2005; Kawasaki, Herrenkohl, & Yeary, 2004; Pnevmatikos, Kariotoglou, & Nikolopoulou, 2006, Smith., Snir, & Grosslight, 1992) was taken into consideration in designing the relevant activities of the TLS. More specifically, students found to consider some irrelevant factors to explain floatation such as (a) the weight of an object, (b) the size of an object, (c) whether or not it had holes or it was hollow, (d) how heavy it was for its size, (e) whether the material had air in it, (f) how much water there was in the container and (g) the kind of material it was made of. Furthermore, some children showed by both their pattern of judgments and their justification that they regarded density as an extensive quantity. More specifically, the activities of the five units were as follows.

Unit 1

In the beginning of the 1st unit of the TLS students are familiarized with f/s phenomena and the relevant concepts, e.g. density, through the study of a technological problem, the salvage of a sunken ship, attempting to find solutions to this problem. This way, on the one hand we enhance the authenticity of the TLS activities and on the other hand we prompt intentional learning (Vosniadou, 2007). A driving question in a contextualized real-world environment is asked: "why did the cruise ship 'Sea-Diamond' sink?" This question serves the purpose of organizing and implementing all the other key questions and activities of the TLS, for example, "Why do ships sometimes sink?", "Which are the variables that affect f/s phenomena?" (Krajcik, 2001). It is known that a carefully designed classroom discussion about the nature and role of models, in parallel with the use of certain models in specific situations, could be significantly helpful for understanding basic characteristics of this kind of epistemological knowledge of science (Vosniadou, 2007). Thus, in the 1st unit a discussion about the nature of models takes place, with the aid of two different (physical, sketch, see Figure 1) models of a ship. In this discussion it is emphasized the fact that these constructions are on the one hand representations of the real cruise ship 'Sea-Diamond' without having to be a replica of it, as well as that they serve as tools in order to explore the reasons and conditions of floating and sinking of the real ship.

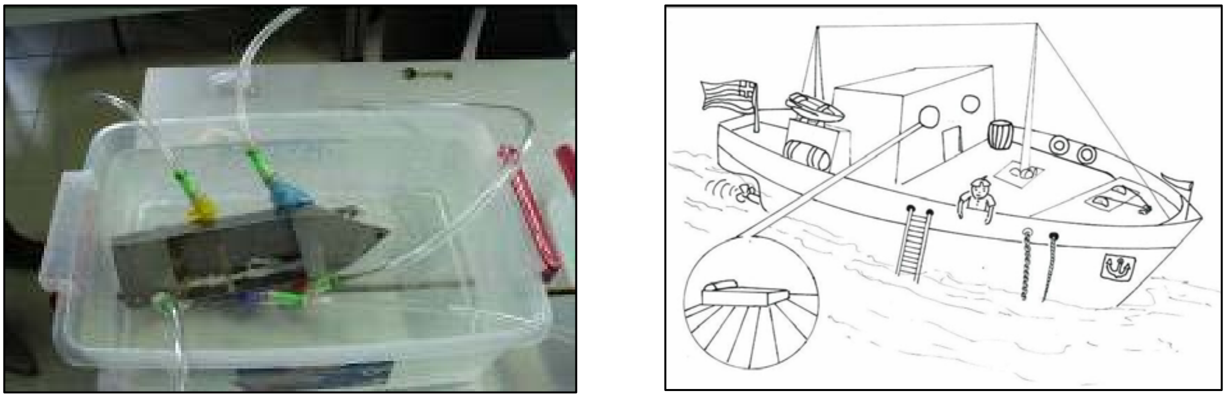
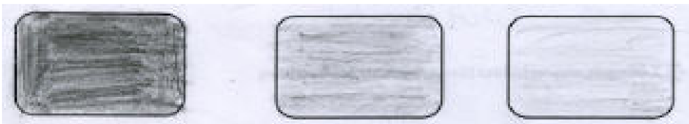


Figure1. The physical and sketch representations of a ship

Unit 2

As a consequence of the initial questions that students are trying to answer in order to hypothesize the reasons of the sinking of the cruise-ship and propose possible solutions in order to salvage the sunken ship, in the 2nd unit children passed in scientific world implementing, in groups, real and simulated experiments in order to identify and control possible factors that affect f/s using CVS method. In this unit there is no explicit external-instructional models' use. However, students' involvement in the above research is important because they encounter and confront their own alternative ideas in reasoning about f /s phenomena.



RUBBER IRON WOOD

Figure 2. Students' 'shade representation' of the 'weighter-lighter' relationship of same volume objects

Unit 3

In the 3rd unit the children are asked to represent, in a paper and pencil task, the 'weighter – lighter' relationship of three objects of the same volume and different material (wood, rubber and iron). Working in a simulated environment, they could use a balance in order to compare the weight of the objects. Children proposed different representations of this relationship, like 'shade representation' (see Figure 2) (see Smith et al. 1992). This is assumed to be an important scaffolding step in order children to be introduced to another visual representation of density, the 'dot crowdedness' model (see Figure 3). The qualitative 'dot crowdedness' model of density (Smith et al., 1992) was adopted instead of a mathematical introduction of the concept, as a consequence of the difficulties children of this age range have to understand mathematical ratios.

Figure 3. The 'dot crowdedness' model of density

Using this representation in relevant simulated experiments children were expected to infer more abstract causal relationships like the predictive rule about f/s. For example, children had the opportunity to experiment in software 'room' 1 (see Figure 4), with two cubes (one made from wood and one from rubber) while letting them in a tank filled with water. They were supposed to conclude a relational predictive law about f/s of objects in water (if the number of dots in the cube of the material of an object is bigger than the number of dots in the cube of water then the object will sink and if the number of dots in the cube of the material of an object is smaller than the number of dots in the cube of water then the object will float).

Summing up the first three units, followed a gradual approach of the nature and the role of models moving from physical models like an iron-made ship representation (1st unit, Figure 1, left), to symbolic models like the visual model of density (3rd unit, Figure 3) and finally to relational models concerning the rules of f/s (3rd unit, "when an object has bigger density than water, the object sinks, and when an object has smaller density than water, the object floats"). This transition is assisted by the negotiation and confrontation of the students' causal (linear) reasoning concerning the factors (e.g. the weight of the object) that affect f/s (2nd unit) so as to promote the passing from a simple linear causality to an interactive relational causality (Perkins & Grotzer, 2005). Significant scaffold in this shift is considered to be the 'weighter-lighter' relationship task, in the beginning of the 3rd unit.

Unit 4

In the 4th unit, two different representations of the heliocentric model (a physical model and a sketch) were used as well as children's visual representations of density, in order to enhance the discussion about models' features. More specifically, it was expected that children would generalize the concept of model in another context except from floating and sinking. Moreover, children would acquire a more integrated view about the nature and the role of models. For example, students could realize that heliocentric model can be expressed in more than one representation (physical and sketch), thus achieving the possibility of multiple representations of models. The heliocentric model's representations were used because children are familiarized with these representations in school and by the media as well.

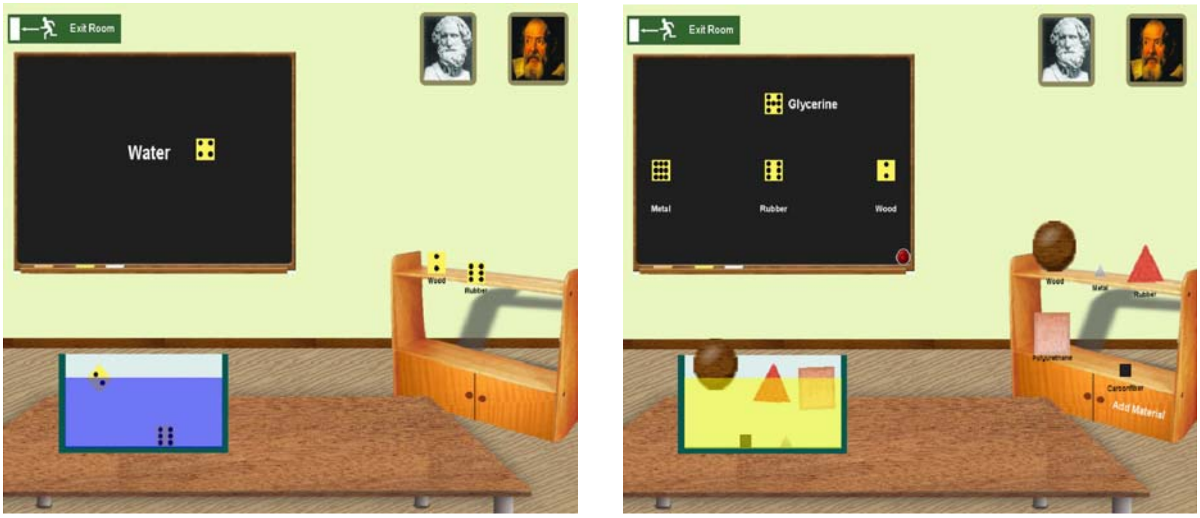


Figure 4: Software 'room' 1 (left) and 2 (right)

In addition, in the 4th unit, in order to help further students' shift from causal linear to causal relational reasoning about f/s, children were asked to experiment in software 'room' 2 (Figure 4). Students had the opportunity to experiment with three objects (a wooden sphere, an iron and a rubber pyramid), letting them in a tank filled with glycerin, in order to generalize the predictive law that they have already concluded, concerning f / s of several objects in water. During this activity, students tested whether these three objects of different shape, size and material float or sink in a tank filled with glycerin. The 'dot crowdedness' model for each one of the materials, including the liquid as well, is available on the blackboard of the software 'room'. The students should be able then to conclude that "if an object's 'dots-per-cube' number, are less than a liquid's 'dots-per-cube' number, of the same volume, then the object will float". Furthermore, children were asked to propose a number of 'dots-per-cube' for polyurethane (the big pink cube which floats) and for carbon-fiber (the black small cube which is sunken) according to the evidence they collect from their experimentation. They were expected to conclude that the number of 'dots-per-cube' for polyurethane should be smaller than glycerin's and the number of 'dots-per-cube' for carbon-fiber should be bigger than glycerin's.

Unit 5

Finally, in the 5th unit, we come back to the technological problem that has been driving all the TLS from the 1st unit, specifically with the question "why did the cruise ship 'Sea-Diamond' sink?" A new technological problem was set then with the question, "how could we salvage the sunken cruise-ship?" Children, initially, had the opportunity to experiment with the salvage of the 'Sea-Diamond' cruise-ship in a simulation (Figure 6). In parallel, with the aid of the simulation they discussed for the last time in the TLS about the nature and role of models, including extensive discussion about the multiple representations of a model, the explanatory and predictive role of a model and the constituents of a model.

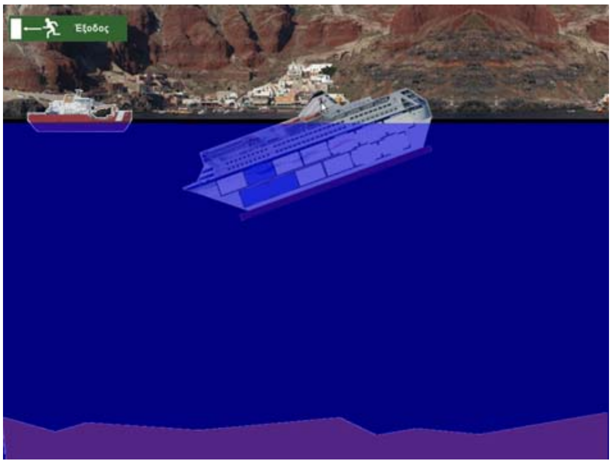


Figure 5: 'Sea-Diamond' simulation

Hereupon, students are trying to find solutions using two real representations of the technological problem. The first one consisted of a physical model of a statue and the second one of an iron – made physical model of a ship, both sunken in a tank filled with water (see Figure 1, left).

Data Collection

Data were collected from multiple sources. However, here, for the purpose of this paper, only the pre and post-questionnaires are presented. Children were asked to write the most representative phrase, according to them, which includes the word model (Q1). Moreover, children were asked to describe the usefulness of a physical model of an eye (Q2), and of two sketches of a ship (Q3) respectively. Additionally, in the post questionnaire only, children were asked to describe the way that the visual representation of density, which they negotiated with during the implementation, has helped them (Q4).

Criteria for responses' assessment were based on (Gilbert, 1991; Spyrtou, Zoupidis, & Kariotoglou, 2008; Treagust et al., 2002). More specifically, it was examined whether children hold beliefs that correspond to the following aspects: (a) a model is a representation of a target, (b) a model is not a copy of the target, (c1) a model (physical, sketch as well as a visual one) helps explain, or predict a phenomenon, (c2) a model (physical, sketch as well as a visual one) can be used in experiments or to understand, (c3) a model (physical, sketch) can be used in experiments or to understand and (d) a model's function is the recreational one (with the meaning of beauty, aesthetics and having fun).

Results

Based on the relevant literature (Treagust et al., 2002, Gilbert, 1991), children's responses were classified into five categories. First, children considered models as a representation that it is not a replica of the target (Q1), while recognized that each of the presented constructions or models help in explaining and/or predicting of a phenomenon in questions (Q2, Q3, & Q4). For example, one student answered in the post questionnaire about the model of density that "... *the model of density has helped us to see whether an object will sink or float* ...". These ideas are considered as reflecting the 'scientific view' for the models. Four children were found holding the scientific view for models after the implementation, while none of them had these ideas before the implementation. Second, children considered models as a representation that it is not a replica or a copy of the target (Q1) and they recognized that each of the presented constructions or models could be used in experiments or to understand phenomena (Q2, Q3, & Q4). For example, one student answered "... *models are representations of real objects but are not exactly the same* ...". These ideas are considered as partially reflecting the scientific view ('Partially scientific view 1'). Twelve children were found holding this partially scientific view for models after the implementation, while none of them had these ideas before the implementation. Third, children considered models as a representation (Q1) without any mention to the target object, and they recognized that each of the presented constructions or models could be used in experiments or to understand phenomena (Q2, Q3, & Q4). These ideas are again considered as partially reflecting the scientific view ('Partially scientific view 2'), but having a bigger distance from the scientific view than the Partially scientific view 1. Eleven children were found holding this partially scientific view for models after the implementation, while none of them had these ideas before the implementation. In other words, more of the children, 27 out of the 41 children, changed their ideas about the models, after the implementation, to the scientific or to the partially scientific. Fourth, children considered models as recreational (Q1), although they recognized that each of the presented constructions or models could be used in experiments or to understand phenomena (Q2, Q3, & Q4). These ideas are considered as alternatives ('Alternative view') to the scientific views, mainly because despite the fact that they can recognize a scientifically accepted function of a model, in parallel they refer to its recreational function as well. Fourteen children were found holding this alternative view for models before the implementation, and ten after the implementation. From the 14 children that held initially the alternative view only three remained at this view, while the rest (11 children) moved to either to the partially or to the scientific view. Finally, children considered models as recreational (Q1), without any answer to the rest of questions about the usefulness of the physical models. These ideas are considered as recreational ('Recreational view') because children recognize only a recreational aspect of the models. For example, a student answered "… *a model is beautiful and very thin* …". Twenty seven children were found holding this alternative view for models before the implementation, and only four after the implementation. That is, after the implementation a great number of children shifted from their simple recreational ideas for models adding at least some other aspects for their use in science.

Conclusions and Implications

The aim of the present study was to examine if it is possible to improve children's epistemological awareness about the role and nature of models, through the gradual introduction of models, from the physical to more abstract. Floating and sinking phenomena and density concept were used as a framework to introduce crucial aspects of the role and nature of models. Children, before the implementation, held recreational and alternative view for the nature and role of models. That is, they claimed that models' function is recreational and none of the participants mentioned the explanatory or the predictive role of models. These results strengthen Gilberts' (1991) evidence that children during the school years hold the recreational view about the role and the nature of models. After the designed implementation, lasting ten teaching hours in five units, most of the children shifted towards the scientific view of models, claiming that a model is a representation and not a copy of the target and referring to the explanatory and predictive role of the models. Only seven, that initially expressed the alternative or recreational view, hold their initial view about the role and nature of models after the implementation. That is, the gradual introduction of models from physical to more abstract ones, as Petrosino (2003) proposed, could indeed facilitate schoolchildren to enhance their epistemological awareness about the role and nature of models. The movement towards the scientific epistemological awareness, however, is a cumbersome procedure and it happens gradually. In agreement with Treagust et al. (2002), children found to attain the descriptive role of the models, and later on the explanatory and predictive role of the models, as well as that model is not replica of the target.

Teaching practices should take into account that the use of models in school settings without a special consideration of the gradual introduction of the models do not assure the epistemological awareness that is needed for deliberate and intentional mechanisms that scientists use for hypothesis testing and conscious belief-revision (Vosniadou, 2007). The gradual introduction of models starting from the physical models to causal relational, through sketch and visual models could facilitate children to understand aspects of models and to intentionally use them to explain and predict phenomena such as floating and sinking. Models are considered as one of the facilitators of conceptual understanding and achievement. Further analysis of the current evidence (see Zoupidis, Spyrtou, Pnevmatikos, & Kariotoglou, in preparation) has shown that in general children that achieved the scientific view of the models, at the same time they grasped the concepts of floating, sinking and density. On the contrary, children, who failed to change their initial view about the nature and the role of models, had difficulties to change their alternative beliefs about the concepts are taught too.

Acknowledgements

Work presented in this paper has been supported by the European Union through the European Communities Research Directorate General in the project Materials Science –University-school partnerships for the design and implementation of research-based ICT-enhanced modules on Material Properties, Science and Society Programme, FP6, SAS6-CT -2006-042942).

We would like to thank the students, teachers and the principle of the 1st primary school in Florina, Greece.

References

Boudreaux, A., Shaffer, P., Heron, P., & McDermott, L. (2008). Student understanding of control of variables: Deciding whether or not a variable influences the behavior of a system. *American Journal of Physics*, *76*(2), 163-170.

Duit, R., 1999, A model of educational Reconstruction – A framework for research and development in Science Education. In: Koumaras, P. et al. (Eds), Proceedings of the 1st Panhellenic conference on Science Education and New Technologies (pp. 30-34). Thessaloniki, Greece: Christodoulides.

Fassoulopoulos, G., Kariotoglou, P., Koumaras, P. (2003). Consistent and Inconsistent Pupil's Reasoning about Intensive Quantities, The Case of Density and Pressure. *Research in Science Education, 33,* 71-87.

Gentner, D. (1983). Structure-Mapping: A Theoretical Framework for Analogy. *Cognitive Science, 7*, 155-170.

Gilbert, S. W. (1991). Model building and a definition of science. *Journal of Research in Science Teaching, 28(*1), 73–79.

Halloun, I., (2004). *Modeling Theory in Science Education*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Havu, S. (2005). Examining young children's conceptual change process in floating and sinking from a social constructivist prospective. *International Journal of Science Education, 27*(3), 259-279.

Kariotoglou, P., Psillos, D., Tselfes, V. (2001). Investigating scientific practice: A means to link research and development. In: D. Psillos et al. (Eds), *Proceedings of the Third International Conference on Science Education Research in the Knowledge Based Society* (pp. 233-235). Thessaloniki, Greece: Art of the Text.

Kawasaki, K., Herrenkohl, L., Yeary, S. (2004). Theory Building and modeling in a sinking and floating unit: a case study of third and fourth grade students' developing epistemologies of science. *International Journal of Science Education, 26*(11), 1299-1324.

Krajcik, J. (2001). Supporting Science Learning in Context: Project-Based Learning. In: R. Tinker, & J. Krajck, (Eds), *Portable Technologies: Science Learning in Context.* New York: Kluwer Academic/Plenum Publishers.

Lijnse, P. L. (1995). Developmental Research As a Way to an Empirically Based Didactical Structure of Science. *Science Education 79*(2), 189-199.

Méheut, M. & Psillos, D. (2004). Teaching-Learning sequences: aims and tools for science education research. *International Journal of Science Education, 26*(5), 515-535.

Nersessian, N. (2008). Model-based reasoning in scientific practice. In Richard A. Duschl and Richard E. Grandy (Eds.), *Teaching Scientific Inquiry: Recommendations for Research and Implementation*, (pp. 57-79). Rotterdam, The Netherlands: Sense Publishers.

Perkins, D., Grotzer, T. (2005). Dimensions of causal understanding: the role of complex causal models in students' understanding of science. *Studies in Science Education, 41*,117-166.

Petrosino, A. (2003). Commentary: A framework for supporting learning and teaching about mathematical and scientific models. *Contemporary Issues in Technology and Teacher Education, 3*(3), 288-299.

Pnevmatikos, D., Kariotoglou, P. and Nikolopoulou, E. (2006). Children's conceptions about floating and sinking: A microgenetic approach. Paper presented at the 5th European symposium conceptual change. Stockholm, Sweden (June 2006).

Smith, C., Snir, J., Grosslight, L. (1992). Using Conceptual Models to Facilitate Conceptual Change: The Case of Weight-Density Differentiation. *Cognition and Instruction, 9*(3), 221-283.

Spyrtou A., Zoupidis A., Kariotoglou P., (2008). The design and development of an ICT-enhanced module concerning density as property of materials applied in floating / sinking phenomena,. Paper presented at the GIREP2008 International Conference, Physics Curriculum Design Development and Validation. Nicosia, Cyprus, (August 2008).

Treagust, D., Chittleborough, G., Mamiala, T. (2002). Student's understanding of the role of scientific models in learning science. *International Journal of Science Education, 24*(4), 357-368.

Vosniadou, S. (2007). The cognitive-situative divide and the problem of conceptual change. *Educational Psychologist, 42*(1), 55-66.

Zoupidis, A., Spyrtou, A., Pnevmatikos, D., & Kariotoglou, P. (in preparation). Does epistemological awareness enhance conceptual knowledge? The case of density in floating and sinking phenomena.

ADAPTIVE LEARNING ENVIRONMENT FOR CHEMICAL REACTIONS

Felix Warbruck & Karin Stachelscheid
University of Duisburg-Essen

Abstract

Adaptive learning environments become popular in educational research. The kind of adaptation varies with different learning environments. In this study we are creating an adaptive learning environment for Grade 9 chemistry courses concerning chemical reactions. We are currently adapting two main factors for learning performance – prior knowledge and learning task difficulty. We adjust the learning tasks with respect to students' ability according to Vygotsky's zone of proximal development (ZPD). The learning tasks are not too easy or too difficult for a student and appropriate to for enhancing learning achievement. Items which fit into the ZPD of a student range between 0.6 logits more or less than student's ability. This can be calculated with the Rasch model. Prior knowledge is one crucial predictor for learning outcomes. For adapting the learning environment to students' prior knowledge, we include test before every chapter and students have to work at chapters they do not know sufficient.

Introduction

Computer-based learning environments became popular for education in the last twenty years. We are developing an adaptive computer learning program, which can be used in everyday school education, where students have significant learning efforts and are interested in the topic of chemical reactions.

The national educational standards for chemistry, introduced in 2005, define students' skills and abilities after Grade 10 and what they are supposed to perform. Yet it is still unknown how to teach students according to the standards. Our learning program is based on the national educational standards for chemistry (KMK, 2005).

Theoretical Background

Design of the learning environment

When creating a new computer-based learning envoronment for chemical reactions, it is necessary to think about the optimal structure of the content. For instance which topic must be understood to get the next topic? How can chemical reactions being described? The content of this learning environment is structured based on research about students' conceptions and misconceptions about chemical reactions (Calık & Ayas, 2005; Schmidt, 1998; Ahtee & Varjola, 1998; Treagust, Chittleborough, Mamiala & Thapelo, 2003; Gilbert & Treagust, 2009, and more) and previous projects for cumulative learning like SINUS.

Chemical reactions can be surveyed in three different representations – macroscopic, symbolic and submicroscopic (figure 1). The macroscopic level describes everything, what can be observed, like material properties. Chemical symbols and reactions equations are the main point at the symbolic level. An adequate construct of matter is necessary for the submicroscopic level. The stages we present in this study include current research about students' conceptions and misconceptions of chemical reactions. The learning program is divided into three stages, first the macroscopic, second the symbolic and third the submicroscopic. In the first stages students should be able to describe the macroscopic perspective of chemical reactions. Students can describe and compare properties of educts and products. In the following stage, students are introduced to the symbolic representation of chemical reactions. They are expected to use molecular formulas for chemical equations and apply the correct terminology (e.g. atom, molecule or element) to describe chemical equations. In the third stage, students gain insight the submicroscopic level of chemical reactions. They understand what happens with atoms and molecules in chemical reaction, e.g. the reorganisation of atoms and the breaking and reforming of bonds. Students have to create appropriate mental representation for these technical terms. This is the last stage of Ahtee & Varjola (1998), who explained it as "sound understanding."

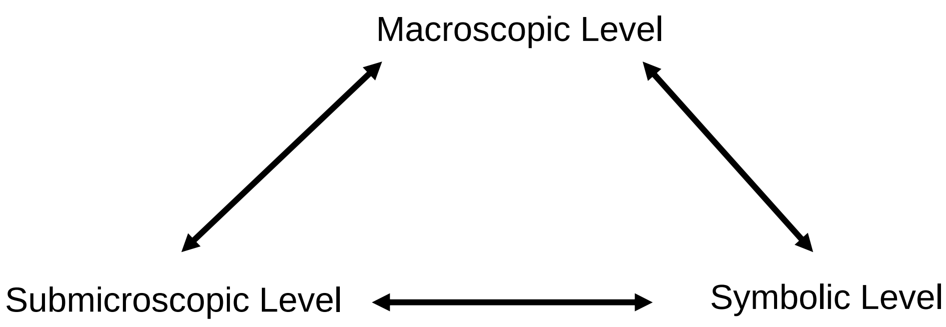


Figure 1. Representations for chemical reactions (e.g. Johnstone, 1982).

A chemical reaction is apprehended if it can be described at all three levels. Example: At the macroscopic level of magnesium combustion flames and smoke can be observed. The educt magnesium is solid and has a silver metallic colour, the product is white and solid. The comparison of educt and product indicates that a chemical reaction has occurred. At the symbolic level, the reaction equation should be arranged ($2\ Mg + O_2 \rightarrow 2\ MgO$). At the submicroscopic stage, students should have an appropriate mental model for this reaction. They should be able to translate the reaction equation to a pictorial representation. Then the student should be able to explain which bonds are broken and that the atoms are rearranged to magnesium oxide. The aim of this learning environment is the description of several chemical reactions at the three different levels of representations. To reach the aim of this study every level of representation is introduced by itself. Later the different levels of representations will be combined.

Adaption of Learning Tasks

Vygotsky's zone of proximal development (Vygotski, 1963) is used as a basic design principle for learning and instruction (Caiklin, 2003) and is already used in some computer based learning programs. We use the idea of the zone of proximal development (ZDP) for learning tasks, which fit to students' current ability. These learning tasks are still in the students' ZPD and they should be able to solve them. We use Rasch analysis to estimate parameters for each person and for each task.

Students with different abilities have different probabilities of solving one item. It could be very easy to solve one item for a high competent student but difficult for a low competent student. We define the ZPD as the individual item solving probability between 20% and 80%. The item difficulty, which fits to students' ability, can be calculated with the mathematical form of the Rasch model. The mathematical form of the Rasch model for dichotomous data includes the probability of correct answer (P), the item difficulty (δ) and the person ability (β) $P = \frac{e^{\beta-\delta}}{1+e^{\beta-\delta}}$. Person and item parameters are measured at one scale with the Rasch model. If classical test theory is used then the item difficulty ranges between 0 and 1 and person's ability is the score of correct solved items. But with Rasch measurement item and person parameters are measured at one scale. If a person has the ability +2, items with difficulty higher than +2 are more difficult than students ability and items lower than +2 are easier than person's ability. In this way items and persons can be compared. Person ability and item difficulty are measured in logit scale. The person ability will be received from a pre-test. When students' ability is known and the probability of correct answer in the Rasch formula is defined, the item difficulty can be calculated. Imagine a student with ability 1 and the solving probability of an item should be 20%, then the calculated item difficulty is 0.4. If the item solving probability for this student should be 80% then the item should have a difficulty of 1.6 logits. The ZDP or the difficulty of items, which fit to students with ability 1 ranges from 0.4 to 1.6. In this way you can calculate the item difficulty which fits to student's ability. If the student's ability is known, the range of the ZPD for item difficulty is 0.6 logits higher or lower than student's ability. So every person will receive learning tasks, which are not too easy or too difficult.

The learning tasks are constructed based on a structure model of competence (figure 2). In this model there are two difficulty generating dimensions. The first one is the complexity (i.e. the linkage level of content elements), which is divided into five levels. The lowest complexity is on the linking level of just one fact, followed by two facts, then one relation and on the fourth level, two relations. The highest level is called "superior concept". The second dimension is the cognitive process. There are four stages: replication, selection, organisation and integration. These two dimensions make up 20 cells. We use the operationalization of the structure model in order to generate learning tasks of a defined difficulty.

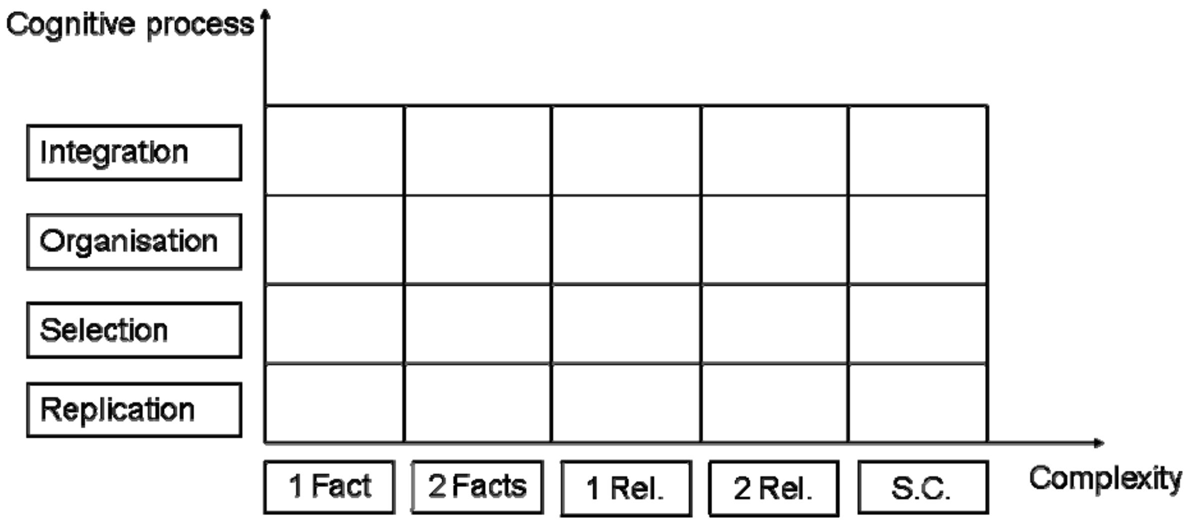


Figure 2. Structure model of competence (Walpuski et al., 2008).

Adaption to prior knowledge

Prior knowledge is one crucial predictor of learning outcomes (Renkl, 1996). For adapting the learning environment to students' prior knowledge, we are including tests before every chapter. Students have to do a test about this chapter and if they fail they have to work at this chapter. If they performed it well, they have to do the next one. Students will receive a feedback of their knowledge. In this way students get to know, what they still do not know sufficient and at which chapters they have to work.

Methods

The study is designed in a 2 x 2 factorial design (figure 3). We will distinguish between adapted and non-adapted learning. "Adapted" learners receive the topics and learning tasks based on tests of their current knowledge. Non-adapted learners can choose their chapters and the learning tasks by themselves.

We are going to recruit classes in Grade 9 that will have already received a chemistry course for one year. The classes will be divided into four treatment groups. The groups are matched by the covariables "general cognitive ability", "reading ability", "self-concept" and "prior knowledge of chemical reactions". Thus, we are able to receive data, which is free of influence due to different school classes. The general cognitive ability is measured by a cognitive ability test by Heller & Perleth (2000). The reading performance is controlled by an ability test (Schlagmüller & Schneider, 2007). For measuring prior knowledge, we are developing an adequate test. Students' interest will be measured by a test which is based on the questionnaire about students' interest (Schiefele, Krapp, Wild & Winteler, 1993). The students have to work about six school lessons with the computer program. The learning efforts and changes of interest will be measured by a pre-post-test and follow-up design.

As mentioned before, we are developing an adaptive computer learning environment for school practice. The research questions of this study are:

1. Do adapted learning tasks lead to higher learning outcomes?

2. Does adaption to prior knowledge perform to higher learning success?

		Adaptive Learning tasks	
		Yes	No
Adaption to Prior knowledge	Adapted	Treatment A	Treatment B
	Individual	Treatment C	Treatment D

Figure 3. Design of the study.

Conclusions and Implications

In this way of teaching we first analyze students' abilities after giving them an adequate learning environment. Students will be promoted individual based on objective tests, which measure their knowledge and their performing with knowledge. Most individual learning environments let students choose their themes or let them choose other things. This is a very subjective choice, which does not base on objective criteria. This adaptive learning environment measures students' abilities objective, so it should enhance learning success more than learning environments, which are based on individual choices.

References

Ahtee, M. and Varjola, I. (1998). Students' understanding of chemical reactions. *International Journal of Science Education*, 20 (3), 305 – 316.

Caiklin, S. (2003). The Zone of Proximal Development in Vygotsky's Analysis of Learning and Instruction. In Kozulin, A., Gindis, B., Ageyev, V.S. & Miller, S. M.(Ed.), *Vygotsky's Educational Theory in Cultural Context* (pp. 39 – 64). New York: Cambridge University Press.

Calık, M.& Ayas, A. (2005). A Comparison of Level of Understanding of Eighth-Grade Students and Science Student Teachers Related to Selected Chemistry Concepts. Journal of Research in Science Teaching, 42 (6), 638 – 667.

Gilbert, J. K. & Treagust, D. F. (Eds.) (2009). *Multiple Representations in Chemical Education.* Dordrecht: Springer.

Heller, K. A. & Perleth, Ch. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen Revision (KFT 4-12+R).* Göttingen: Hogrefe.

Johnstone, A. H. (1982). Macro- and micro-chemistry. *School Science Review*, 64, 377–379.

Renkl, A. (1996). Vorwissen und Schulleistung. In J. Möller & O. Köller (Hrsg.), *Emotionen, Kognitionen und Schulleistung* (S. 175-190). Weinheim: Beltz.

Schiefele, U., Krapp, A., Wild, K.-P. & Winteler, A. (1993). Der "Fragebogen zum Studieninteresse" (FSI). *Diagnostica*, 39, 335 – 351.

Schlagmüller, M. & Schneider, W. (2007). *Lesegeschwindigkeits- und verständnistest für die Klassenstufen 6-12*. Göttingen: Hogrefe.

Schmidt, H.-J. (1998): Students' Misconceptions - Looking for a Pattern. *Science Education*, 81, 123 – 135.

KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (2005). *Beschlüsse der Kultusministerkonferenz – Bildungsstandards im Fach Chemie für den mittleren Schulabschluss (Jahrgangsstufe 10).* München: Luchterhand.

Treagust, D. F., Chittleborough, G. & Mamiala, T. L. (2003). The role of submicroscopic and symbolic representations in chemical explanations, *International Journal of Science Education*, 25 (11), 1353 – 1368.

Vygotski, L. S. (1963). Learning and mental development at school age (J. Simon, Trans.). In B. Simon & J. Simon (Ed.), *Educational psychology in the U.S.S.R.* (pp. 21 – 34). London: Routledge & Kegan Paul.

Walpuski, M., Kampa, N., Kauertz, A. & Wellnitz, N. (2008). Evaluation der Bildungsstandards in den Naturwissenschaften. *Mathematisch Naturwissenschaftlicher Unterricht*, 61, 323 – 325.

A METHODOLOGY TO ASSESS CONCEPTUAL PROFILES WITHIN CONTEXT DELIMITATION

Cristiano Mattos & Esdras Viggiano
University of São Paulo

Abstract

The conceptual profile model (Mortimer, 1995) conceives that an individual possesses several mental representations of a same concept, which are used in different contexts. In this work we propose a questionnaire as instrument to assess the pre-service Brazilian Physics teachers' conceptual profiles of "to teach" and "to learn". The questionnaire was developed privileging the connection between conceptual profile zones and theirs context of use. We propose, in certain way, a more general method to assess conceptual profiles. As result, we present an outline of categories and dimensions that allows, in certain degree, the control of the contextualization in each question. Thus, during the application of the questionnaire we intended to drive the respondents to contexts that interest to us. In this work we presented the preliminary results that show the validity of the questionnaire.

Introduction

Nowadays one of the major consensus in education is that individuals' prior ideas influence the way they think and learn. A prolific Science Teaching research had been established to study these priors' ideas in the 1970´s and 1980´s. A lot of research effort was developed to propose methods to change prior ideas to the scientific ones. After a couple of decades, Science Teaching research pointed out that those changes, as replacement of the old to the scientific ones, were limited indicating a cohabitation of different ideas. To deal with those results Mortimer (1995) proposed the Conceptual Profile model. It is composed by several cognitive representations of a single concept – called of Conceptual Profile Zone (CPZ) – that are used according with the context.

In spite of being an important presupposition of the Mortimer's (1995) conceptual profile model proposed the role of the contexts was rarely worked out. We identify in Silva's (2006) and Zaine's (2003) works concerns about how the zones of the conceptual profile are used in contexts. Nevertheless, the concern about contexts becomes visible only during data analysis and not as presupposition when producing the assessment instruments (Viggiano 2008).

An important characteristic of the conceptual profile model is the evolution of conceptual profile (Mortimer, 1995). It consists in the aggregation of new zones of conceptual profiles, e.g. scientific zones. It means that the conceptual profile changes including new zones, which are used in different contexts. From this point of view students continue to use prior zones in some contexts thinking that are most adequate than others. A possible perspective to Science Education is to delimit the contexts of use of scientific concepts pointing out the most appropriate zones. Mattos and Rodrigues (2007) proposed a hierarchical learning scheme to deal with conceptual profile dynamic, point out "learning orders" to explain the learning process as a hyper-contextualization process instead a de-contextualization one.

Rationale

The notion of context, besides indispensable, has a central function in the process of construction of the assess instrument. We used Bernstein´s (1990) recontextualization theory that refers to attribution of meanings in contexts as bases of our model. Here we propose that the conceptual profile zones are used according context defining specific meanings, i.e., the same principles and rules that regulate context shaping and meaning making also regulate the zones of conceptual profile. Bernstein (1990) says that contexts are established or defined in the social interaction, and they are governed by the locational communicative principle and interactive communicative principle. The locational principle attributes spatial sense to contexts and the interactional principle establishes the temporality of the interactions among subjects. Furthermore, there are context recognition rules which are the main responsible for its delimitation, establishing the contextual borders, supplying marks that distinguish a context of another. The attribution of meanings occurs in contexts defined by the human communicative interaction and in agreement with the predominance of the communicative principles and the realization rules. These rules are established using internal contextual features, linked with the textual elements (deixis), with the communicative principles and with the others contextual rules. For Bernstein (1990), several contexts can coexist in the same local. It´s occur because the contexts are established in communication, and more than one can happen at the same time. Besides, one general context (macro-context) can include more specific contexts (micro-contexts).

The same principles and rules that regulate context recognition are the same ones that structure the use of zones of conceptual profile in each context. This principles and rules socio-historically define context hierarchical structures we model as micro and macro-contexts. Our intention is point out a unit of analysis respecting hierarchical influences that could define some of the context's markers used to recognize meanings in specific contexts (Mattos and Rodrigues, 2007)

An illustrative example (Figure 1) is the educational system. In the classroom, teacher interacts with a student, establishing a micro-context, a group of student interacts with other and the teacher creating other micro-context, and the Physics Class delimits a macro-context including these micro-contexts. However, when we look at the school as a whole macro-context, other classes such as chemistry, biology etc., as well the physics classes are micro-contexts in the school macro-context.

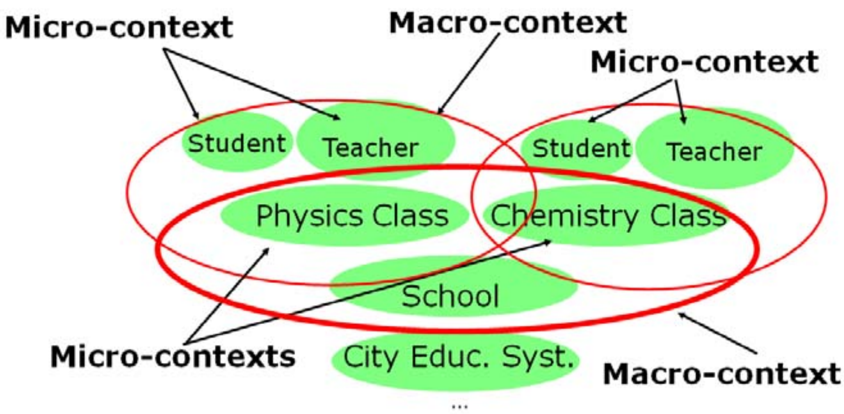


Figure 1. A propose of multiply contexts architecture.

Taking in account the physics teacher training, we constituted our sample with students of pre-service physics teacher of University of São Paulo (Brazil). We assumed that the pre-service teachers´ uses of their mental representations according to contexts influences how they teach. So when we identify and present how they conceptualize some aspects of the teaching-learning process, we might produce reflections about their practice and about their own formation course. However, simply showing their representations do not necessarily implies they realize new ways to represent or change their performance at classroom. The knowledge about a specific educational concept is limited considering the diversity of possible contexts of use if we consider the conceptual profile as a whole. Nevertheless, present to pre-service teacher their conceptual profile related with their profession should point out the ways he/she perform as a teacher in different contexts allowing them integrate practical knowledge with other aspects of his/her life.

Hence to know about the use of conceptual profiles zones of educational related concepts, as *learning*, *teaching*, *teacher*, *student*, *school*, *to teach*, *to learn*, *education*, *evaluating*, *curriculum* etc. in different contexts contributes to teacher's formation courses. These conceptual profiles influence significantly how the teachers work in the classroom.

Within the framework presented, we propose to build a questionnaire to assess conceptual profiles of *to learn* and *to teach*. We believe that the methodology used can be used with other conceptual profiles.

To build the questionnaire it was necessary to explore deeper the relations between the zones of the conceptual profile and specific contexts where they are used.

Methods

Despite the relevance of contexts in conceptual profile model, these have not been considered in the most researches focused in conceptual profiles assessment instruments. Stressing the importance of contexts, we propose in this work a questionnaire, assuming a narrow connection between context and conceptual profile zones. This instrument was structured from a priori categories schema, picked up from literature of Science Education Research to define some contexts of use of the conceptual profiles zones of "to teach" and "to learn" (Mattos & Viggiano, 2008). These categories are the base to organize and select textual elements (*deixis*), which establish context recognition rules. Then the questions intent to provoke the emergence of realization rules, that driving subjects to use specific conceptual profile zones of "to teach" and "to learn". The category group was structured as a net, with four hierarchies, as represented in the figure 1.

We considered the conceptual profile as the cognitive representation of a complex system of contexts. Thus, the conceptual profile itself is a cut of that complexity, also presenting characteristics of a complex system. The figure 2 intends to represent a cut in that complexity of contexts and, consequently, serving as a set of criteria to access its projection at the subjects' conceptual profiles.

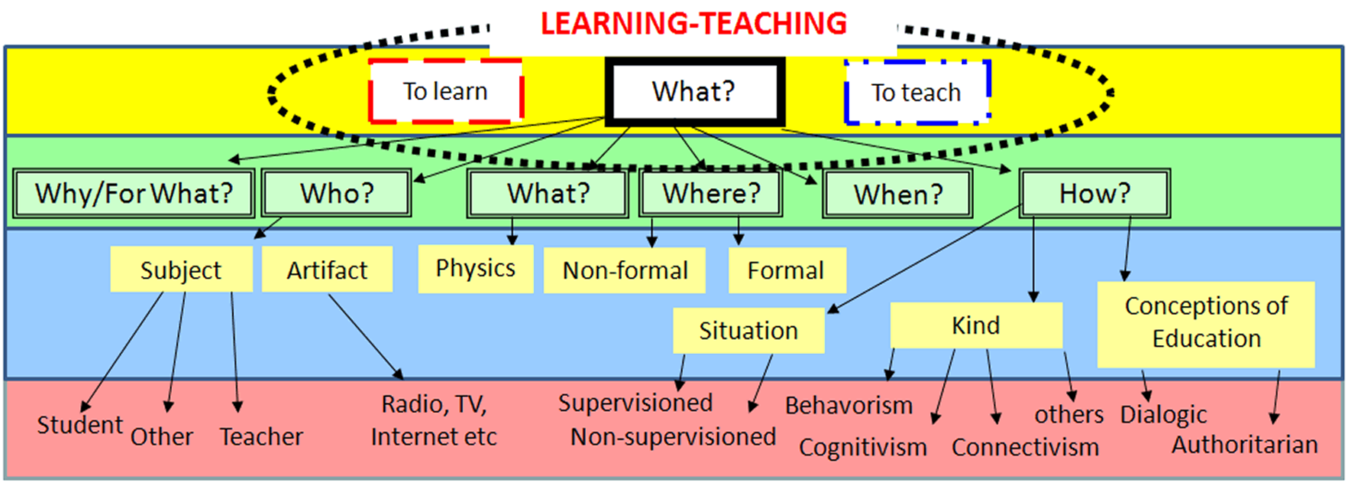


Figure 2. Scheme of categories associated with the conceptual profiles of "to teach" and "to learn".

We propose that a conceptual profile include different hierarchical levels that could be identified as others conceptual profiles (Viggiano, 2009). Assumed as a complex system a conceptual profile allows us to think about the relationship between multiple conceptual profiles, which together constitute a more complex one. Some conceptual profiles have a sharper zone containing no other simpler conceptual profiles. For instance, it is possible to conceive that conceptual profiles of "to teach" and "to learn" can be subordinated to a superior hierarchy that would represent the conceptual profile of "to teach-to learn". We identified signs of that profile, since we observed correlations among the use of "to teach" and "to learn" zones in similar contexts (Viggiano, 2008).

The tree structure as showed in Figure 2 represents *a priori* categories scheme. There are four hierarchies. The first one is the most general and serves as origin to the lower ones and refers to the category "What is [to teach/to learn]?".

Mattos and Rodrigues (2007) and Dalri and Mattos (2008) extended the conceptual profile notion including an axiological dimension that refers to values and ends of concepts. This dimension influences how individuals use and constitute theirs conceptual profiles. Considering this idea, we built the second hierarchy with the category "For what?". This category, take in account the axiological dimension, dealing with information about the beliefs and values related with the teaching-learning process.

Another category of the second hierarchy corresponds to the question "What [to teach/to learn]?" related with the specific content (in our case Physics) and its relations with the teaching-learning process.

At the same hierarchy is the question "Who?" that refers to who is involved in teaching-learning process. A third hierarchical level corresponds to the sub-category of "Who?" referring to the "Person" who learns or teaches, the other sub-category is the "Artifact", constituted by the fourth hierarchical categories: TV, book, radio, internet, etc.

Another second hierarchy category is referred to "Where?" and refers to where occurs the teaching-learning process, bringing information about the sub-categories formal and non-formal education.

The last category of the second hierarchical level corresponds to the question "How?" that have three subcategories at third hierarchical level. The first dimension of third order is the "Situation" of teaching and learning that could be understood as a situation where there are or not a person whom supervise or regulate this kind of situation. The sub-categories (forth hierarchical level) indicate this division by the "Supervised" or "Non-supervised" categories.

The second category of "How?" (third hierarchical level) is "Conceptions of Education" which refers to how occur the interactions between the subjects evolved in educational process. It has the categories of four order "Authoritarian" and "Dialogic". Another dimension of third order is "Kind" and refers to characteristics related to theories of development and learning of psychology.

The questionnaire was composed with questions done in pairs. This means that a question about teaching was also done to learning. The questions were built from terms referring to the categories of our schema. Our intension was to define the contexts in a way that answers should manifest the most probable zone used in this context by the individual.

The question "what is [to teach/to learn]?" leads the hierarchy of context control. When we went down in the hierarchies more delimited were contexts that questions refers.

The questions were built to contemplate all the hierarchical levels presented in the Figure 2. Our expectation was that questions with more defined context will provoke bigger chances of use specific conceptual profile zones. At same time less defined the context more the variety of the conceptual profile zones will emerge.

Contextualization, in the sense of context delimitation, is given, by the question, with the specification of *what, why/for what, who*, *how* and *where* "to teach" and "to learn". We tried to control the context in two ways.

The first one, we use top-down categories defining sharper contexts when going deeper in the hierarchical levels and, in some way, sending the subject to a specific contexts. In other words higher the hierarchy less specific is the context implying lesser rules of recognition due the limited semantic *deixis* (or language parameters) offer. For instance, "How is possible [to teach / to learn]?" and "What is [to teach / to learn]?" are in the first hierarchy and are less contextualized than "How to put yourself in the others´ position to help you to teach them?" built in four hierarchy.

The other way we used cross-categories questions, with more than one categories and hierarchical levels. In the both cases were introduced specific communicative marks of context (deixis) using some words related with priori categories from the schema. For instance, the question "Where does the teacher teach out of school?" links the following categories: *Non-formal* with *Who/What/Person/Teacher*.

Almost whole questionnaire was built using this method. Only some associations were established due the uncountable possible options and time spend by students to answer an extensive questionnaire.

As we said before, every question was created in pairs, e.g. "What is 'to learn'?" and "What is 'to teach'?", allowing us to compare answers of the same category and context with different concepts.

After the application of a pilot questionnaire, we attain to 44 questions, distributed in four sub-questionnaires as showed at Figure 3.

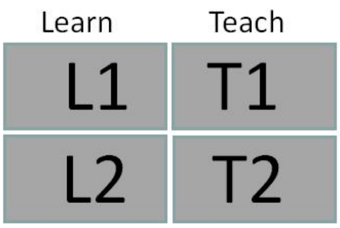


Figure 3. Organization of questions.

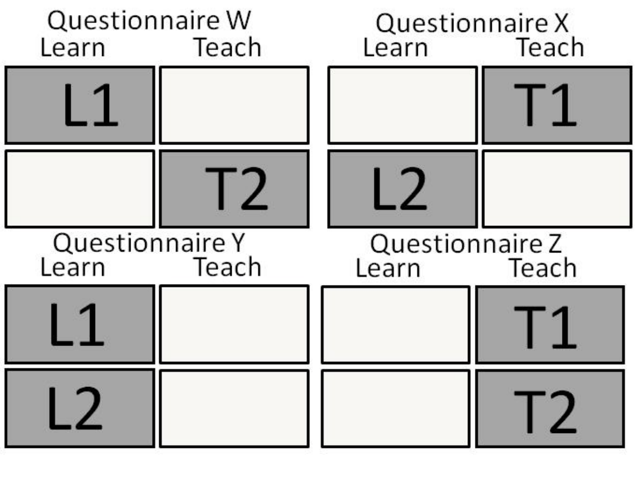


Figure 4. Structure of the four sub-questionnaires.

Figure 3 shows the "L" blocks, which correspond to 22 questions about "to learn". The "T" blocks are the other 22 questions about "to teach". "L" questions are each one paired to another "T" questions (L1 with T1, and L2 with T2).

We wanted to investigate if the sequence of questions could induce answers, and at same time validate sub-questionnaires as possible candidates to larger assessment. We distributed the questions into the four sub-questionnaires as shown in Figure 4.

Our sample was composed by 40 Brazilian pre-service Physics teachers of the teacher formation course at University of São Paulo (Brazil), each one answered one of the four models of questionnaires.

Analysis

In this work we categorized answers tanking account expressions that reflect the categories of two dimensions – *Where* and *Conceptions of Education*, each one with three sub-categories. For example, to the questions "What is 'to teach'?" we had answers as "It´s to accumulate knowledge." that had been categorized as *Authoritarian*, and answers as "It´s to share knowledge with colleagues" categorized as *Dialogic*. The main ideas referred by each category are showed in Table 1.

Table 1. Main ideas that of *Conception of Education* category.

Category	"to teach"	"to learn"
Authoritarian	Instructor transmits knowledge actively.	Learner receives knowledge passively.
Dialogic	The knowledge is shared actively by learner and teacher. Teaching-learning process is meaningful to both teacher and student	The knowledge is shared actively by learner and teacher. Teaching-learning process is meaningful to both teacher and student
Mixed	The answer doesn't have sufficient information to be classified.	The answer doesn't have sufficient information to be classified.

Authoritarian category was characterized by terms representing authoritarian speech, a simultaneously inconsiderateness of teacher and learner in educational relation, or not considering teachers and students' mutual learning. "To teach" is considered as an active contends transmission and "to learn" as passive contends reception. Moreover, non-consideration of individual characteristics contributes to categorize answers in this category.

Answers in *Authoritarian* category can also be characterized by expressions that reveal inconsiderateness of students' previous knowledge. The typical expressions representing teaching were verbs as *to introduce*; *to transmit*; *to pass*; *to show*; *to do* and *to supply*. The learner assumes the role of a passive knowledge receiver in the educational process. The typical expressions representing learning were also verbs as *to absorb*; *to acquire*; *to assimilate*; *to receive*; *to take knowledge*; *to adapt, to receive* and *to accumulate*.

On the other hand, answers categorized as *Dialogic* consider students' needing and the importance of previous knowledge. Education becomes a process of knowledge meaning share between teacher and students, resulting in mutual teaching and learning.

Differently from category *Authoritarian*, the typical expressions used in *Dialogic* category were not verbs but sentences. To express "to teach" were used some expressions as *to change experiences*; *to provide conditions for another acquires knowledge*; *help other to develop his owner knowledge*; *allows that the other learns for himself*; *to stimulate the student*; *to share knowledge*, etc. There are lot of sentences to express "to learn" in this category. Therefore, some answers were long and complexes ones demanding to understand the whole idea to categorize it.

The answers in the *Mixed* category refer to mixed or dubious answers. They did not contain enough elements to be classified in one of others categories or show characteristics of the both ones. An example of a frequent answer in this category is "to learn is to expand the knowledge".

We show in Table 2 some examples of categorization of answers to the question "What is to learn?".

Table 2. Examples of categorization on *Conception of Education* categories.

St	Q01L – What is to learn?	Category
001	Learn is make knowledge stay with you.	Mixed
002	To learn is not objective, nor rational. It´s subjective and emotional. It´s not accumulation of data or knowledge. It´s discuss. Even if you don´t obtain answers. To learn is make questions (it´s not answers them). I don´t believe in Absolute True.	Dialogic
008	Learning does not occur only at school, studding regular contents, the learning occurs since the moment that we born	Mixed
032	It´s to acquire knowledge about one or several subjects, explaining facts or situations.	Authoritarian

The answer of Student 001 was categorized as doubtful because he only considers "to learn" as a product and not as a process. The verb "to stay" refers to knowledge, the product of learning process. He did not give enough information about how the process happens, it does not allow us to distingue between authoritarian and dialogic categories. Then the answer had been categorized as mixed or doubtful.

The Student 002 said that learning "is not accumulation of data or knowledge", rejecting the idea of an "Absolute true". When he said "to collect data or knowledge", showed implicitly that disagree with the vision of learning as "accumulation of knowledge". In this sense, he may know that there are several ways to conceive teaching and learning. The Student 002 refers to subjective and emotional characteristics of that learning, indicating the Dialogic category.

The Student 008 does not allow us to identify what "to learn" means to him. Referring to the dimension *where*, he made clear that learning occurs out of school (non-formal context). The response did not give enough information to be categorized as Dialogic or Authoritarian.

The Student 032 relates the knowledge learned with its use. When used the word "I", he highlight that is how is his own learning.

We categorized the answers in *Where* dimension at the same way, looking to formal or non-formal places. To make the categorization of answers, we looked for words that reveal the consideration of formal spaces, for example, the words *school*, *teacher*, *student*, *class*, *laboratory etc.*

Table 3. Pearson´s correlation (*r*) among questionnaires considering the categories *Conceptions of Education (CE) and Where (W).*

Questionnaire	*CE´s r*	*W´s r*
W and Y (L1)	+ 0,764	+ 0,902
W and Z (T2)	+ 0,729	+ 0,858
X and Y (L2)	+ 0,831	+ 0,828
X and Z (T1)	+ 0,793	+ 0,821

* All values were significant for $\alpha \leq 0.01$.

We build four tables of frequencies using pair of questions from different sub-questionnaires (W and Y; W and Z; X and Y; X and Z). We used Pearson's Linear Correlation Test (Muijs, 2004) on data presenting the Pearson´s correlation coefficient (*r*) at table 1. The results could be considered significant in according the method used ($\alpha \leq 0.01$). The correlations were significant, and their values are close to +1, in a scale [–1,+1], meaning that higher values on one variable goes together with higher values on another variable".

In our research, we consider *r* as positive strong when ($r > 0.5$) and very strong when ($r > 0.8$), indicating that independent of questionnaire model, the answers were almost categorized in the same categories.

Results

The descriptive levels obtained indicate that the placement of questions on any of the questionnaires leads to similar results. Therefore, we can choose any of the questionnaires to apply in a larger sample. Moreover, we interpreted these descriptive levels as an indicative that the own questions are enough to define the contexts of use of conceptual profile zones. Furthermore, these indicate that working with pairs of questions built from a priori categories schema we can obtain similar uses of conceptual profile zones in near contexts, regardless of grouping questions. Hence, the questions were sufficient to establish the recognition and framing contextual rules, implying in the resemble use of conceptual profile zones.

Conclusions and Implications

The correlation obtained indicates that each pair of questions was enough to define contexts, evoking the use of similar conceptual profile zones. This correlation also indicates the possibility to assess conceptual profiles with a smaller number of questions. Mainly because different questionnaires had significant descriptive levels among themselves, i.e., at any group of questions we could identify the same conceptual profile zones, indicating that the questions drive subjects to the same or similar contexts.

The a priori categories schema seemed adequate to build instruments of assessment of conceptual profiles considering the context of use of conceptual profile zones. We believe that this kind of scheme can be used to assess others conceptual profiles, giving the context his appropriate importance in the communication processes.

References

Bernstein, B. (1990) **The Structuring of pedagogic discourse**: Volume IV Class, Codes and Control. London: Routledge & Kegan Paul, 1990.

Dalri, J.; Mattos, C.R. (2008) Relaciones entre motivación, valor y perfil conceptual: un ejemplo. In: Noveno Simposio de Investigación en Educación en Física, 2008, Rosario. Memorias del … Rosario: Asociación de los profesores de la Facultad de Ciencias e Ingeniería de la Universidad Nacional de Rosario.

Mattos, C.R.; Rodrigues, A.M. (2007) Theoretical Consideration about conceptual profile dynamic. In: ESERA Conference, 2007, Malmö. **Proceedings ESERA Conference**. Malmö: Malmö University, 1. 1-12.

Mattos, C.R.; Viggiano, E. (2008) An instrument for assessment of the conceptual profiles of to Teach and to Learn. In: **Proceedings of GIREP-EPEC Conference**. Rijeka: Faculty of Arts and Sciences - University of Rijeka2008.

Mortimer, E. F. (1995) Conceptual Change or Conceptual Profile Change? **Science & Education**, 4, 267-285.

Muijs, D. (2004) **Doing quantitative research in education with SPSS**. London: SAGE.

Silva, F.A.R. (2006) **O perfil conceitual de vida: ampliando as ferramentas metodológicas para sua investigação.** Master´s Dissertation. UFMG (Brazil).

Viggiano, E. (2008) **A proposal of conceptual profiles assessment of ´to learn´ and ´to teach´.** Master´s dissertation. Physics Institute/Education Faculty of University of São Paulo.

Viggiano, E. (2009). Perfil conceitual complexo: a interação entre perfis conceituais. **Enseñanza de las Ciencias**, Extra, pp. 3068-3073.

Zaïne, N. (2003) Conceptual profile of pupils and students of the radiation. In: ESERA Conference. **Proceedings ESERA Conference**, Noordwijkerhout, Netherlands.

ENHANCING COGNITIVE SKILLS IN SCIENCE BY PROVIDING COGNITIVE TUTORING WITH AUTHORING TOOLS: A CASE STUDY OF A COLLABORATIVE SOFTWARE ENVIRONMENT IN A GREEK PRIMARY SCHOOL

Theodoros Panagiotopoulos
University of Aegean

Evi Makri Botsari
School of Technological and Pedagogical Education, ASPETE

Sarantos Psycharis
School of Technological and Pedagogical Education, ASPETE

Anastasia Reppa
School of Technological and Pedagogical Education, ASPETE

Abstract

This study concerns the use of a cognitive tutoring tool to primary school students within a collaborative software environment. Our teaching proposal is based on the assumption that we can both improve metacognitive and learning skills of students using an intelligent tutoring system (ITS), the CTAT, by adding collaborative learning elements to this software. The process followed two steps: First we intergrated Cool Modes, a collaborative software tool, within CTAT. Second we developed educational-didactic scenario in which 3 dyads (6/12) students used the software in order to collaborate in solving problem-based modelling tasks. The system was evaluated in a real classroom, and the results showed that students' metacognitive skills and learning performance increased significantly.

Introduction

ITS are highly interactive learning environments that have been shown to improve typical classroom instruction (Murray 1999, VanLehn 2006) ,they have been successful in raising student achievement (Koedinger et.al 2003) and have been disseminated widely. ITS are based on artificial Intelligence technology to provide interactive instruction that adapts to individual students' needs and personality's characteristics and, most typically, support student's practice in learning through complex problem solving and reasoning. Cognitive Tutors are a type of intelligent tutor based on cognitive psychology theory of problem solving and learning. Cognitive Tutors also provide a rich problem-solving environment with tutorial guidance in the form of step-by-step feedback, specific messages in response to common errors, and on demand instructional hints. They also select problems based on individual student performance (Koedinger & Aleven 2007) and grew out of an attempt to apply and test the ACT-R theory of cognition and rely on cognitive psychology research in their design and development (Anderson, 1993). Studies have found some evidence to the connection between students' metacognitive decisions, while working with ITS, and their learning performance (Aleven et al. 2003, Baker et al. 2004, Wood and Wood 1999).

Rationale

In the last decade, research on metacognitive guidance embedded within collaborative environments has been under research, and its positive effects on students' achievement, reasoning, and discourse have been documented (e.g. Mevarech & Kramarski, 2003; Veenman et al., 2006). Recent research efforts have indicated metacognition as one of the strategies that can improve learning performance and various efforts have been made to extend Cognitive Tutors to support different kinds of meta-cognitive and reflective thinking processes in collaborative environments (e.g., Schoenfeld, 1992; White & Frederiksen, 1998).

Indicatively we refer:
1. Self-explanation (Aleven & Koedinger, 2002)
2. Error self-correction (Mathan & Koedinger, 2005)
3. Avoiding "gaming" the instruction (Baker et al., 2006)
4. Help-seeking skills (Roll et al., 2007).

In our work, we are interested in using Cognitive Tutors combined with the collaborative environment provided by Cool Modes. We assume that computer-mediated collaboration can promote metacognitive skills and conceptual knowledge in physics. In our approach, called bootstrapping novice data (BND), we provided students with a computer-based tool, let them attempt to solve problems with the tool, and record that problem-solving activity in a tutor-specific representation. BND implementation is realized through the integration of a collaborative modeling tool, called Cool Modes (Collaborative Open Learning and MODEling System) (Pinkwart 2003), and a tutor authoring environment/Intelligent Tutoring System, the Cognitive Tutor Authoring Tools (CTAT) (Koedinger et al. 2004).

Cool Modes was developed to support "conversations" and shared graphical modeling facilities between collaborative learners (Pinkwart, 2003). This is achieved through a shared workspace environment with synchronized visual representations. These representations together with their underlying semantics can be defined externally by plug-in visual languages and interpretation patterns, so-called "reference frames". This is achieved through a shared workspace environment with synchronized visual representations. (McLaren, Koedinger, Schneider, Harrer, & Bollen, 2004).

Cognitive Tutors Authoring Tools (CTAT) is a type of intelligent tutor based on cognitive psychology theory of problem solving and learning. CTAT supports the creation of both Cognitive Tutors (Koedinger & Corbett, 2006) and a newer type of tutors called example-tracing tutors. Example-tracing tutors evaluate student behavior by flexibly comparing it against examples of correct and incorrect problem-solving behaviors. Example-tracing tutors are capable of sophisticated tutoring behaviors: they provide step-by-step guidance on complex problems while recognizing multiple student strategies and maintaining multiple interpretations of student behavior. They can be built without programming, through drag-and-drop techniques and programming by demonstration. Example-tracing tutors have been used in real educational settings for a wide range of application areas (Aleven 2006).

The goal of our paper is to evaluate if the metacognitive support through CTAT related with the collaborative Cool Modes would help children to benefit during collaboration. The different tutorial scenario were built in a context where, unlike earlier tutorial dialogue systems, students have the freedom to explore and communicate to each other. We've implemented a scenario, based on the subject of electric circuit, into a real classroom (5^{th} and 6^{th} grade students), collected the data from log files and analyzed the metacognitive indicators and the learning performance.

Methods

Authoring an example-tracing tutor for CTAT includes four steps:
1. Creation by the expert of a graphical user interface
2. Forecast and demonstration of correct, incorrect and alternative solutions in the problem
3. Annotation of the steps of solution in "behavior recorder" with hint messages, feedback messages and headings that are related with actions, concepts or skills.

The approach of solving problems with CTAT is quite tedious. The major problem was to predict the possible solutions, the correct and incorrect actions of students according to the services provided by CTAT. In case –aw in our case- which we want to merge CTAT with collaboration, there are many possible sequences and alternative paths, so it was hard to interpret and compare the students' actions in the Behavior Graph. Additionally, some groups increase the state space by annotating a problem or by expressing equal sequences of actions of different users. The logs of the different collaborating groups should be translated into a single behavior graph in the Behavior Recorder. After the behavior graph is generated, a tutor author manually updates it by adding hints and bug messages, annotating buggy paths, and adding skills to the edges (McLaren et al, 2004).

We applied the CTAT and Cool Modes methodology in children of for the course of Physics (Grades 5, 6). The computer based learning lessons were conducted in a real classroom. The school was experimental and we had the chance to have a lot of degrees of freedom about some changes in the curriculum. Twelve computers – one for each student –had been provided to us from our school exactly for this the purpose. The participants were picked randomly, 6 females and 6 males participated. All the subjects were assigned to dyads (1 male and 1 female). For all of them it was the first time that they used both the Cool Modes and CTAT, so we trained them for six months for the use of the software.

The aim of our research was to compare the metacognitive development and learning performance between the experimental –called also metacognitive group(A) and the control group(B). Group A used simultaneously CTAT and Cool Modes (Group A) and Group B only the CTAT.

There were three dyads per group (six in each group A,B), and the methodological procedure was the standard pre-test/intervention/post-test. The students completed pre- and post- tests before and after the study. Tests followed a logical, detailed process of dynamically generated material presentations, hints, feedback and monitoring (covered or open), and assessment of declarative knowledge using hypothetical metacognitive assistant tutoring. The pretest and posttest problems were equivalent.

The students collaborated working on separate computers and they were free on using a number of tools that scripted and supported their collaboration within Cool Modes. In the metacognitive mode (group A), the tutor provided metacognitive support via hints sent to the students to promote explanations, reflection, and help giving/receiving.

The tutor used the Behavior Recorder to observe and recognize situations requiring a hint before the provision of that. Examples of metacognitive hints were "Don't forget to explain your statements and actions to each other," which was provided by CTAT when a student neglected to explain a statement or action despite a request by his partner, and "Remember to talk about and reach a consensus on your next activity before moving on" which was provided when a student started an activity alone before agreeing on previous activities with his/her partner.

The instruction with CTAT, using various representative tools, presupposes the creation of educational scenario. The domain selected for implementing our educational scenario was based on electricity and the relationship among relationship among potential difference, current and resistance as well as issues related to charged objects. We asked students to construct a diagram for a single cell, light bulb and switch to be placed together in a circuit such that the switch can be opened and closed to turn the light bulb on and secondly to solve problems about electricity, which are representative of typical problems. And to answer questions about e.g. the luminosity of bulbs in a circuit, he value of current for the different bulbs etc.

The problems were equivalent for both groups and equivalent (not similar) before and after the intervention. During problem solving steps, both groups had access to simple forms hints and simple kind of feedback (e.g. try again) but the experimental group(Group A) was provided by metacognitive support through Cool Modes.

Cool Modes uses a visual language which employs text cards for the items (e.g. voltage sources, resistors), numerical indicators for ammeters (for measuring current), and links to connect items to the wires of the circuit. Since Cool Modes uses a consistent, internal naming scheme for objects, we can leverage this to identify common dynamic objects across sessions using mapping tables. Behavior Recorder then collects information about each component (type and values), the way students' connect wires and measure quantities used in the circuit along with the nodes to which they are connected. For example, to solve one of the problems one dyad of the experimental group requires to change the values of circuit voltage and note the corresponding changes in the current keeping the resistance constant. For such circuits, Cool Modes allows the student to use the range of component values indicated in the workspace. If the student uses values of components other than those specified in the workspace the BR identifies that as an error.

Quite often a circuit will not work for the first time. We offered hints on an incorrect entry or hints on the next problem-solving step. The first, or top-level hint, is very general and abstract (e.g. try again). If a circuit did not work, we sent a general message to the students telling them to check the battery or to check if the wires are properly connected. The purpose is to remind the student what action to take. If students ask for a second hint, this hint is considered as more specific and direct. For the experimental group the ultimate goal is to make the students able to create good reasoning for taking a step.

Finally, the last hint, or the *bottom-out hint*, tells the student explicitly what action to take. After the intervention students had to solve problems, as we've planned in the post test design. The "metacognitive support group" with the aid of tutoring system(CTAT) and the Cool Modes was prompted after each step to explain why the step is logical, what is its role in solving the problem, how it relates to what they know already, etc. The given task was to discover the equations that allow them to do these calculations.

Results

Studying metacognition one faces the problem of using valid criteria-techniques-instruments that measure students' metacognitive development. The complexity of this task arises from two main sources: the lack of a generally accepted conceptualisation of what really metacognition means and second due to the fact that metacognition is an inner awareness or process rather than an explicit behaviour and consequently individuals themselves are often not conscious of these processes (Georgiades, 2004).

In our work we used the CTAT environment in order to code the behaviour of teacher students. In CTAT once an interface is created, the author can use it and the associated "Behavior Recorder" to create problems and demonstrate alternate solutions. We suggest that CTAT can measure metacognition due to the nature of the behaviour recorder to register all the strategies followed by the students as well as the kinds of additional information he/she asks during problem solving.

CTAT also provides the student with a rich problem-solving environment with a variety of representational tools and presents authentic problem scenario that require reasoning, as well as additional tools, such as a worksheet, a symbolic equation solver e.t.c. for the student to solve. For certain errors that students commonly make, CTAT presents an error feedback message that explains why the step is wrong and in addition to error feedback messages, at any point in the problem scenario, students can request help from the tutor. CTAT also presents hints that are specific to the solution strategy taken by the student and multiple levels of hints are available which explain how problem-solving principle can be applied.

According to the above mentioned aspects of CTAT we decided to use the characteristics of CTAT for measurements of metacognition, and we focused on two components of that namely: the knowledge about cognition and regulation of cognition.

Metacognition measured before the intervention using the following indicators for 10 problems that students had to solve.

- self-evaluating their ideas,
- self-questioning when they encountered blocks,
- detecting their errors, considering a range of possible alternatives,
- and considering limitations in their ideas (Linder & Marshall 1997).

Each of the above mentioned indicators were marked as 1(for success) and 0 (for no action of students). Aspects of metacognition activity that were evaluated during problem solving in the environment of CTAT(for both groups) included the following three indicators.

1. **Students strategies** (at each step of each problem-1 grade for each reasonable strategy). Strategy was recognized as such by the instructor if it contained techniques to solve problems
2. **The number of efforts to find alternative solutions** (1 grade for each alternative solution). Students could solve for example the problems using kinematics relation or energy theorems
3. **The number of efforts to solve the problem** (1 grade for each reasonable effort).

Table 1. Mean Metacognition of Group A

GROUP A – USE OF CTAT/COOL MODES	Mean metacognition	N
before the intervention with CTAT	60	6
after the intervention with CTAT	87	6

Table 2. Mean Metacognition of Group B

GROUP B – USE OF CTAT	Mean metacognition	N
before the intervention with CTAT	58	6
after the intervention with CTAT	75	6

Table 3. Mean learning performance of Group A

GROUP A – USE OF CTAT/COOL MODES	Mean learning performance	N
before the intervention with CTAT	45	6
after the intervention with CTAT	78	6

Table 4. Mean learning performance of Group B

GROUP B – USE OF CTAT	Mean learning performance	N
before the intervention with CTAT	43	6
after the intervention with CTAT	70	6

Low metacognition corresponds to the range 0-30/100, medium 30-70/100 and high 70-100/100. The classification was carried out following an analysis of the log files provided by CTAT.

A cognitive task analysis of log data recordings taken during the experiments showed that the control-group dyads were less likely to correct flaws in their collaborative and script practice. On the other hand, the hints given to the metacognitive dyads, although not always appreciated by the students, had a clear positive effect on collaboration, motivation, and to some extent on the way the collaborating students followed the procedure.

Conclusions and Implications

In a primary Greek school six of twelve students received metacognitive support using the Cool Modes software in connection with CTAT.. The data presented indicate that we can promote effective learning in physics if we design metacognitive tasks in a collaborative environment. In our study, the metacognitive support (comprehension, connection, strategy) was based on the integration of Cool Modes and CTAT. We have been using Behavior Recorder, a tool for building Pseudo Tutors, a special type of cognitive tutor which is based on the idea of recording problem solving behaviour by demonstration and then tutoring students using the captured model as a basis. Our method called Bootstrapping Novice Data, involves the transformation of student log files generated by a problem-solving software tool into a sequence of student action messages «evaluated» by the tutor authoring software. Log data provides further evidence that metacognitive support in a collaborative environment has the potential to contribute to metacognitive growth of students than stand alone problem-solving practice. Results have shown that students have improved their domain knowledge relative to control students. We are in the process for future study to expand the investigation of cognitive tutoring approach to collaborative technology environments and to investigate if this implementation is capable to enhance metacognition and foster pedagogical knowledge through problem solving tasks.

References

Anderson, J. R. (1993). Rules of the Mind. Hillsdale, NJ: Erlbaum.

Aleven, V., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive Science*, *26,* 147-79.

Aleven, V., Stahl, E., Schworm, S., Fischer, F., & Wallace, R.M. (2003). Help Seeking and Help Design in Interactive Learning Environments. *Review of Educational Research*, *73*(2), 277-320.

Aleven, V. (2006). An intelligent learning environment for case-based argumentation. *Technology, Instruction, Cognition, and Learning 4*(2), 191-241.

Baker, R. S., Corbett, A. T., Koedinger, K. R., & Wagner, A. Z. (2004). Off-task behavior in the Cognitive Tutor classroom: When students "game the system." In ACM CHI 2004: *Computer–Human Interaction* (pp.383–390).

Baker, R., Corbett, A., Koedinger, K. R., Evenson, S., Roll, I., Wagner, A., Naim, M., Raspat, J., Baker, D., & Beck, J. (2006). Adapting to when students game an intelligent tutoring system. In M.Ikeda, K. D. Ashley, T.-W. Chan (Eds.) *Proceedings of the 8th International Conference on Intelligent Tutoring Systems*, (pp. 392-401). Berlin: Springer-Verlag.

Bunt, A., Conati, C., & Muldner, K. (2004). Scaffolding self-explanation to improve learning in exploratory learning environments. In J. C. Lester, R. M. Vicario, & F. Paraguaçu (Eds.) *Proceedings of Seventh International Conference on Intelligent Tutoring Systems*, ITS 2004 (656-667). Berlin: Springer Verlag.

Carver, S. M. (2001), Cognition and instruction: Enriching the laboratory school experience of children, teachers, parents, and undergraduates. In S. M. Carver, & D. Klahr (Eds.), *Cognition and instruction: Twenty-five years of progress* (385–426). Mahwah, NJ: Erlbaum.

Conati, C. & VanLehn, K. (2000). Toward computer-based support of meta-cognitive skills: A computational framework to coach self-explanation. *International Journal of Artificial Intelligence in Education*, *11*, 398-415.

Graesser, R. C., & McNamara, D. (2005). Scaffolding deep comprehension strategies through AutoTutor and iSTART. *Educational Psychologist, 40*, 225–234.

Georgiades, P. (2004). From the general to the situated: three decades of metacognition., *International Journal of Science Education, vol. 26*, 3, pp. 365-383.

Harrer, A., McLaren, B.M., Walker, E., Bollen, L., & Sewall, J. (2005). Collaboration and Cognitive Tutoring: Integration, Empirical Results, and Future Directions. In the *Proceedings of the 12th International Conference on Artificial Intelligence and Education (AIED-05).* Amsterdam, the Netherlands, July 2005. (pp. 266-273). IOS Press.

Harrer, A., McLaren, B.M., Walker, E., Bollen, L., & Sewall, J. (2006). Creating Cognitive Tutors for Collaborative Learning: Steps Toward Realization. User Modeling and User-Adapted Interaction: *The Journal of Personalization Research (UMUAI), 16*, pp. 175-209.

Koedinger, K.R., Aleven, V., & Heffernan, N. (2003). Toward a rapid development environment for Cognitive Tutors. In U. Hoppe, F. Verdejo, & J. Kay (eds.), *Artificial Intelligence in Education*, Proc. of AI-ED 2003. IOS Press, Amsterdam.

Koedinger, K. R., Aleven, V., Heffernan, N., McLaren, B. M., & Hockenberry, M. (2004) Opening the Door to Non-Programmers: Authoring Intelligent Tutor Behavior by Demonstration. In J.C. Lester, R.M. Vicari, & F. Parguacu (Eds.) *Proceedings of the 7th International Conference on Intelligent Tutoring Systems*, 162-174. Berlin: Springer-Verlag.

Koedinger ,K.R & Aleven, V. (2007). Exploring the Assistance Dilemma in Experiments with Cognitive Tutors. *Educational Psychology Review*, *19*.

Koedinger, K.R., A. Corbett, (2006) Cognitive Tutors: Technology Bringing Learning Sciences to the Classroom, in *The Cambridge Handbook of the Learning Sciences*, R.K. Sawyer. p.p 61-78. Cambridge University Press: New York, NY

Kramarski, B., & Mevarech, Z. R. (2003). Enhancing mathematical reasoning in the classroom: Effects of cooperative learning and metacognitive training. *American Educational Research Journal*, *40*(1), 281-310.

Kramarski & Gutman, (2006). How can self-regulated learning be supported in mathematical e-learning environments? *Journal of Computer Assisted Learning*, *22*, 24-33.

Lin, X., Hmelo, C., Kinzer, C.K., & Secules, T.J. (1999). Designing technology to support reflection. *Educational Technology Research and Development*, *47*(3), 43-62.

Luckin, R. & Hammerton, L. (2002). Getting to know me: Helping learners understand their own learning needs through meta-cognitive scaffolding. In S. A. Cerri, G. Gouardères & F. Paraguaçu (Eds.) *Proceedings of Sixth International Conference on Intelligent Tutoring Systems*, ITS 2002 (pp. 759-771). Berlin: Springer.

Mathan, S. A., & Koedinger, K. R. (2005). Fostering the intelligent novice: learning from errors with metacognitive tutoring. *Educational Psychologist*, *40*(4), 257-65.

Mevarech, Z. R., & Kramarski, B. (2003). The effects of worked-out examples vs. meta-cognitive training on students' mathematical reasoning. *British Journal of Educational Psychology*, 73, 449-471.

McLaren, B. M., Koedinger, K. R., Schneider, M., Harrer, A., and Bollen, L. (2004). Towards Cognitive Tutoring in a Collaborative, Web-based environment. In M. Matera, S. Comai (Eds.), *Engineering Advanced Web Applications*: Proceedings of Workshops in *Connection with the 4th International Conference on Web Engineering* (pp. 167-179). Princeton: Rinton Press

Mitrovic, A. (2003). Supporting Self-Explanation in a Data Normalization Tutor. In V. Aleven, U. Hoppe, J. Kay, R. Mizoguchi, H. Pain, F. Verdejo & K. Yacef (Eds.) *Supplemental Proceedings of the 11th International Conference on Artificial Intelligence in Education*, AIED2003., University of Sydney. pp. 565-577.

Murray, T.(1999): Authoring intelligent tutoring systems: An analysis of the state of the art. *International Journal of Artificial Intelligence in Education*, 10.

Nathan, M., Koedinger, K., and Alibali, M. (2001). Expert blind spot: when content knowledge eclipses pedagogical content knowledge. In L. Chen et al. (Eds.), *Proceedings of the Third International Conference on Cognitive Science.* (pp. 644-648). Beijing, China: USTC Press.

Pinkwart, N. (2003) A Plug-In Architecture for Graph Based Collaborative Modeling Systems. In U. Hoppe, F. Verdejo & J. Kay (eds.): *Proceedings of the 11th Conference on Artificial Intelligence in Education*, (pp.535-536). IOS Press

Quintana, C., Zhang,M., & Krajcik, J. (2005). Scaffolded software environments for supporting metacognitive aspects of online inquiry. *Educational Psychologist*, *40*, 235–244.

Roll, I., Aleven, V., McLaren, B., & Koedinger, K. (2007). Can help seeking be tutored? Searching for the secret sauce of metacognitive tutoring. In R. Luckin, K. Koedinger, & J. Greer (Eds.), *Proceedings of the 13th International Conference on Artificial Intelligence in Education* (pp. 203-210). Amsterdam, the Netherlands: IOS Press.

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sensemaking in mathematics. In D. Grouws (Ed), *Handbook of research on mathematics teaching and learning* (pp. 334–370). New York: MacMillan.

White, B. Y., & Frederiksen, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction*, *16*(1), 3–118.

Wood, H. A., & Wood, D. J. (1999). Help seeking, learning, and contingent tutoring. *Computers and Education*, *33*(2), 153–169.

White, B. Y., & Frederiksen, J. R. (1990). Casual model progressions as a foundation for intelligent learning environments. *Artificial Intelligence*, *42*, 99-157.

VanLehn, K. (2006). The behavior of tutoring systems. *International Journal of Artificial Intelligence in Education*, *16*.

Veenman, M. V. J., Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning*, *1,* 3-14.

CONTEMPORARY SCIENCE EDUCATION RESEARCH:

LEARNING AND ASSESSMENT

This book includes a collection of papers on the following topics:

Learning Science
Theories, strategies, and models of learning; meaningful learning, conceptual development, conceptual change, developing understandings, developing skills. Cognitive, affective, and social factors in learning science.

Assessment of Student Learning and Development
Development, validation and use of standardized test, achievement tests, high stakes tests, and instruments for measuring attitudes, interests, beliefs, self-efficacy, science process skills, conceptual understandings, etc; authentic assessment, formative assessment, summative assessment; approaches to assessment.

ISBN - 978-605-364-033-2