

# An evaluation of argument mapping as a method of enhancing critical thinking performance in e-learning environments

Christopher P. Dwyer · Michael J. Hogan · Ian Stewart

Received: 21 April 2011 / Accepted: 12 October 2012 /  
Published online: 27 October 2012  
© Springer Science+Business Media New York 2012

**Abstract** The current research examined the effects of a critical thinking (CT) e-learning course taught through argument mapping (AM) on measures of CT ability. Seventy-four undergraduate psychology students were allocated to either an AM-infused CT e-learning course or a no instruction control group and were tested both before and after an 8-week intervention period on CT ability using the Halpern Critical Thinking Assessment. Results revealed that participation in the AM-infused CT course significantly enhanced overall CT ability and all CT sub-scale abilities from pre- to post-testing and that post-test performance was positively correlated with motivation towards learning and dispositional need for cognition. In addition, AM-infused CT course participants exhibited a significantly larger gain in both overall CT and in argument analysis (a CT subscale) than controls. There were no effects of training on either motivation for learning or need for cognition. However, both the latter variables were correlated with CT ability at post-testing. Results are discussed in light of research and theory on the best practices of providing CT instruction through argument mapping and e-learning environments.

**Keywords** Argument mapping · Critical thinking · e-Learning · Disposition · Cognitive load

## Introduction

Critical thinking (CT) is a metacognitive process that focuses on “purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological or contextual considerations upon which that judgment is based.” (Facione 1990, p. 3). According to Boekaerts and Simons (1993), Brown (1987) and Ku and Ho (2010b), effective learning and problem-solving often involve the application of metacognitive skills, including CT and reflective judgment skills (see also Dawson 2008). CT is made up of a collection of sub-skills (i.e. analysis, evaluation, and inference) that, when used appropriately, increases the chances of producing a logical solution to a problem or a valid conclusion to an argument

---

C. P. Dwyer (✉) · M. J. Hogan · I. Stewart  
School of Psychology, NUI, Galway, Ireland  
e-mail: c.dwyer2@nuigalway.ie

(Facione 1990); and are metacognitive in the sense that they involve the ability to think about thinking (Dawson 2008; Flavell 1979; Ku and Ho 2010b). While cognition is often used to refer to mental processes associated with thinking, according to Flavell (1979), metacognition refers to the knowledge and thinking concerning these cognitive processes and their products. By this definition, critical thinking is largely metacognitive in nature as it involves the ability to analyse and evaluate one's own cognition or the cognition of others and infer reasonable conclusions from the associated cognitive and metacognitive products of thinking in this context.

The teaching of CT skills in higher education has been identified as an area that needs to be explored and developed (Association of American Colleges and Universities 2005; Australian Council for Educational Research 2002; Higher Education Quality Control 1996) as such skills allow students to gain a more complex understanding of the information being presented to them (Halpern 2003). Not only are CT skills important in the academic domain, but also in social and interpersonal contexts where adequate decision-making and problem-solving are necessary on a daily basis (Ku 2009). Good critical thinkers are more likely to get better grades and are often more employable as well (Holmes and Clizbe 1997; National Academy of Sciences 2005).

Previous research suggests that CT can be enhanced through the intervention of semester-long training courses in CT (Gadzella 1996; Hitchcock 2003; Reed and Kromrey 2001; Solon 2007). More specifically, recent meta-analyses suggest that CT can be enhanced by academic courses that directly teach CT or have CT infused in the course (Alvarez-Ortiz 2007), provided that the instruction makes the teaching of critical thinking explicit to students (Abrami et al. 2008). Though research suggests that CT training can improve CT ability, currently, the teaching of critical thinking (CT) remains a veritable challenge (Kuhn 1991; Willingham 2007) for both educators and university students alike. For educators, there is often difficulty in implementing an efficient and effective strategy that targets the teaching and development of CT skills. The challenge for students is overcoming the difficulty associated with simultaneously assimilating and critically thinking about text-based arguments.

Harrell (2005) notes that students often fail to understand the 'gist' (Kintsch and van Dijk 1978) of text-based information presented to them; and more often, students cannot adequately 'follow' the argument within a text (i.e. the chain of reasoning and the justification of claims in the chain), as most students do not even acknowledge that the deliberations of an author within a text represents an argument and instead read it as if it were a story. Another reason why students may find it difficult to assimilate text-based argumentation is that texts often present students with verbose 'maze-like' arguments that consist of massive amounts of text (Monk 2001). Given that text-based arguments contain many more sentences than just the propositions that are part of the argument; these sentences may obscure the intention of the piece and the inferential structure of the argument (Harrell 2004). Though comprehension of an argument's structure may often be difficult for students, abstracting the argument structure from long passages is often necessary, including situations where the argument is highly complex (Harrell 2004; Kintsch and van Dijk 1978). The failure to comprehend the structure of text-based arguments makes the task of critical thinking much harder. Thus, it must be a goal of educators to provide students with the training necessary to analyse and evaluate both simple and complex argument structures.

### The problematic nature of text-based argumentation

Text is presented in a linear fashion, yet text-based arguments are not necessarily sequential and may contain a substantial quantity of verbiage that is not part of the argument. As a

result, one may need to switch attention from one paragraph or page to another and back again in order to assimilate the information within the text (van Gelder 2003). This switching of attention is a cause of cognitive load, which impedes learning by placing added burden on cognitive resources, such as consuming limited working memory space; thus making less space available for the assimilation of the argument (Sweller 1988, 1999). For example, research conducted by Tindall-Ford et al. (1997) found that learning is impeded when instructional materials require a high degree of attention switching. They concluded that encoding environments that increase the cognitive load placed on the reader tend not only to slow the learning process, but also reduce overall levels of learning. Presenting information in a way that reduces the level of attention switching may minimize cognitive load, enhance thinking and improve learning. One such way of presenting information is through argument mapping.

Argument mapping

In an argument map (see Fig. 1), a text-based argument is visually represented using a ‘box-and-arrow’ style flow-chart wherein the boxes are used to highlight propositions and the arrows are used to highlight the inferential relationships that link the propositions together (van Gelder 2003). Specifically, an arrow between two propositions is used to indicate that one is evidence for or against another. Similarly, colour can be used in argument mapping (AM) to distinguish evidence for a claim from evidence against a claim (i.e. green represents a support and red represents an objection). As such, AM is designed in such a way that if one proposition is evidence for another, the two will be appropriately juxtaposed (van Gelder 2001); and the link explained via a relational cue, such as *because*, *but* and *however*. These AM features have been hypothesized to facilitate metacognitive acts of critical thinking, both by making the structure of the argument open to deliberation and assessment; and by revealing strengths and weaknesses in the credibility, relevance, and logical soundness of arguments in the argument structure.

Computer-based argument mapping (AM) is a relatively recent learning strategy (van Gelder and Rizzo 2001; van Gelder 2007) and as such, there is as yet little research examining its efficacy. Nevertheless, available research has identified the use of AM as a strategy that may enhance overall levels of critical thinking (e.g. Alvarez-Ortiz 2007; Butchart et al. 2009; Twardy 2004; van Gelder et al. 2004; van Gelder and Rizzo 2001; van Gelder 2001). For example, in a meta-analysis conducted by Alvarez-Ortiz (2007), it was found that students who participated in semester-long CT courses that used at least some AM within the course achieved gains in CT ability with an effect size of .68 SD, CI

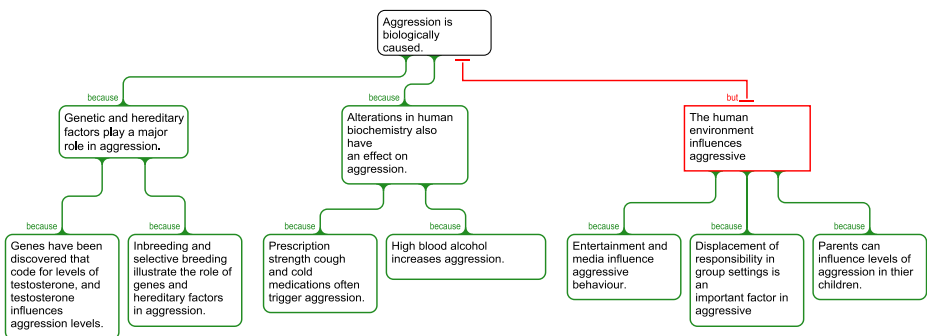


Fig. 1 An example of an argument mapping created through rationale™

(.51, .86). In courses where there was *lots of argument mapping practice* (LAMP) there was also a significant gain in students' CT performance, with an effect size of .78 SD, CI [.67, .89]. Though previous studies of argument mapping have often reported positive effects on critical thinking, however, no firm conclusions concerning the efficacy of this technique can be drawn on the basis of many of these results, because the studies involved have suffered from design limitations. These include the lack of a control or comparison group (e.g. Twardy 2004; van Gelder 2001; van Gelder et al. 2004); not adequately matching or randomly assigning conditions (e.g. Butchart et al. 2009; van Gelder and Rizzo 2001); and the lack of statistical comparison between experimental and control groups (e.g. Butchart et al. 2009; van Gelder and Rizzo 2001). Thus, it is difficult to draw any conclusions about the merit of AM from the studies to date, given the limitations in study design.

More recently, our own research (Dwyer et al. 2011) compared the CT performance of participants on an AM-infused CT course with that of participants on a CT course without AM (i.e. in which instruction utilised traditional means of presentation, such as slideshows consisting of bullet points and outlines, referred to as a 'traditional' CT training course) and a control group. Results indicated that AM training enhanced specific CT sub-skills of evaluation and inductive reasoning. However, this study was also limited in certain respects. Due to a large attrition rate and relatively small resulting sample, the power of the statistical analysis was somewhat diminished.

We suggested that, in order to decrease the problem of attrition, future studies might utilise online courses instead, given that a key limitation in our own research was students dropping out due to competing commitments whereas online courses allow improved time-flexibility. We also minimized instructor feedback in an effort to avoid potential confounds associated with experimenter bias in the comparison between AM training and traditional CT training. However, another potential advantage of e-learning environments is that feedback can be provided in a standardized manner that may circumvent problems associated with experimenter or instructor bias, and feedback itself may be critical for significant learning gains to be observed in CT training studies (Butchart et al. 2009; van Gelder 2003). For example, a meta-analysis of the effects of different methods of instruction on learning, by Marzano (1998) indicated that, across a number of pedagogical studies, feedback concerning (i) strategy used to improve learning and (ii) the efficacy of the use of that strategy produced significant gains in student achievement ( $d=1.31$ ). In addition to the general benefits of feedback on learning, with specific regards to CT, it may be that feedback on CT performance can provide valuable opportunities to evaluate and reflect upon one's own thinking.

### The current research

AM has been developed with the explicit intention to lessen cognitive load and facilitate both the learning and the cultivation of CT skills (van Gelder and Rizzo 2001; van Gelder 2003). First, unlike standard text, AMs represent arguments through dual modalities (visual-spatial/diagrammatic and verbal/propositional), thus facilitating the latent information processing capacity of individual learners. Second, AMs utilise Gestalt grouping principles that facilitate the organisation of information in working memory and long-term memory, which in turn facilitates ongoing CT processes. Third, AMs present information in a hierarchical manner which also facilitates the organisation of information in working memory and long-term memory for purposes of enhancing and promoting CT. In relation to the first reason, dual-coding theory and research (Paivio 1971, 1986), Mayer's (1997) conceptualisation and empirical analysis of multimedia learning, and Sweller and colleagues' research on cognitive

load (Sweller 2010), suggests that learning can be enhanced and cognitive load decreased by the presentation of information in a visual-verbal dual-modality format (e.g. diagram and text), provided that both visual and verbal forms of representation are adequately integrated (i.e. to avoid attention-switching demands). Given that AMs support dual-coding of information in working memory via integration of text into a diagrammatic representation, cognitive resources previously devoted to translating prose-based arguments into a coherent, organised and integrated representation are 'freed up' and can be used to facilitate deeper encoding of arguments in AMs, which in turn facilitates CT (van Gelder 2003).

The second related reason for why AM is hypothesised to enhance overall learning is that AM also makes use of Gestalt grouping principles. Research suggests that when to-be-learned items are grouped according to Gestalt cues, such as proximity and similarity, they are better stored in visual working memory (Woodman et al. 2003; Jiang et al. 2000). For example, Jiang et al. (2000) found that when the spatial organisation, or relational grouping cues denoting organisation (i.e. similar colour, close proximity) are absent, working memory performance is worse, and that when multiple spatial organisation cues (such as colour and location) are used, performance is better. These findings suggest that visually-based information in working memory is not represented independently, but in relation to other pieces of presented information; and that the relational properties of visual and spatial information are critical drivers of successful working memory and subsequently, CT (Halpern 2003; Maybery et al. 1986). Given that related propositions within an AM are located close to one another, the spatial arrangement complies with the Gestalt grouping principle of proximity.

In addition, AM adopts a consistent colour scheme in order to highlight propositions that support (green box) or refute (red box) the central claim, thus complying with the Gestalt grouping principle of similarity (i.e. greens are grouped based on similarity, as are reds). Collating items according to grouping cues, such as similarity (i.e. green because and red buts) and spatial proximity (Farrand et al. 2002; Jiang et al. 2000), may simplify the method of representing information and increase the capacity of visual working memory.

The third reason for why AM is hypothesised to enhance CT is because it presents information in a hierarchical manner. When arguing from a central claim, one may present any number of argument levels which need to be adequately represented for the argument to be properly conveyed. For example, an argument that provides a (1) support for a (2) support for a (3) support for a (4) claim has four levels in its hierarchical structure. More complex or 'deeper' arguments (e.g. with three or more argument levels beneath a central claim) are difficult to represent in text due to its linear nature; and yet it is essential that these complex argument structures are understood by a student if their goal is to analyse or evaluate the argument; and to infer their own conclusions. On the other hand, the hierarchical nature of AM allows the reader to choose and follow a specific branch of the argument in which each individual proposition is integrated with other relevant propositions in terms of their inferential relationship.

The current research examined the effect of AM on CT performance in an e-learning environment in comparison with a no intervention (i.e. neither argument mapping nor e-learning) control condition. Based on the proposed representational properties of AM described above, as well as previous research by van Gelder (2001) and Butchart et al. (2009), we hypothesised that AM training through an e-learning CT course would significantly enhance CT performance. Notably, a number of researchers in the field of CT argue that CT is a domain-general process, in that it can be taught in any educational setting and applied to any academic subject area (Abrami et al. 2008; Ennis 1998; Halpern 2003). For example, a recent meta-analysis conducted by Abrami et al. (2008) investigated the effects

of different CT instruction methods, using Ennis' (1989) typology of four CT course types (i.e. general, infusion, immersion and mixed).

In the general approach to CT training, actual CT skills and dispositions "are learning objectives, without specific subject matter content" (Abrami et al. 2008, p. 1105). The infusion of CT into a course requires specific subject matter content upon which CT skills are practiced; however, this can be any subject area. In the infusion approach, the objective of teaching CT within the course content is made explicit. In the immersion approach, like the infusion approach, specific course content upon which critical thinking skills are practiced is required. However, CT objectives in the immersed approach are not made explicit. Finally, in the mixed approach, critical thinking is taught independently of the specific subject matter content of the course. Results of the meta-analysis revealed that courses that made the CT element of instruction explicit to students (i.e. general, infusion and mixed) yielded the largest CT improvements.

The results of the meta-analysis by Abrami and colleagues indicated that it is the way in which CT is taught (rather than the subject area content upon which CT skills are applied) that is the crucial element in aiding students development of CT skills. Therefore, making CT objectives and requirements clear to students plays an important role in the development of CT ability. In light of research findings indicating the domain-generality of CT and given that participants in the current research were 1st Year Arts students, it was decided that CT course content should cover a variety of topics across academic domains within the Arts programme (e.g. aggression in society, the challenge of work-life balance, attractiveness and social preferences, etc.) in order to: (1) avoid alienating any specific group or groups of students; and (2) maintain the interest of all students (regardless of academic field).

The current research also examined the effects of level of engagement in AM training on CT performance. Previous research by van Gelder et al. (2004) found that CT performance and AM practice hours were significantly correlated ( $r=.31$ ). Therefore, we hypothesised that students who engaged more with the CT course (as measured using the number of AM exercises they completed) would perform significantly better on CT performance than those who did not engage as much.

The current research also examined the relationship between disposition towards thinking and CT ability. A growing body of research has highlighted the importance of this relationship (e.g. Ennis 1998; Halpern 2003, 2006; Ku and Ho 2010a, b; Dwyer et al. 2011). According to Valenzuela et al. (2011), while some conceptualisations of disposition towards thinking focus on the attitudinal and intellectual habits of thinking, many others emphasise the motivational features associated with a positive disposition towards CT. That is, these motivation-focused conceptualisations emphasise the importance of motivation as a process used to activate the metacognitive resources necessary to conduct good CT (Ennis 1996; Norris 1994; Perkins et al. 1993; Valenzuela et al. 2011).

Though few empirical studies have examined the motivational aspects of CT dispositions, research by Valenzuela et al. (2011) revealed that motivation to think critically is a more significant correlate of CT ability ( $r=.50$ ) than is a general positive disposition toward critical thinking ( $r=.20$ ). Similarly, research by Garcia et al. (1992) found a significant, positive correlation between CT ability and motivation towards intrinsic goal orientation ( $r=.57$ ), elaboration ( $r=.64$ ) and metacognitive self-regulation ( $r=.64$ )—three sub-scales of the Motivated Strategies towards Learning Questionnaire (Pintrich et al. 1991). In addition, research has also shown that motivation to learn positively influences CT and learning in general (Hattie et al. 1996; Robbins et al. 2004). Hence, the current research sought to clarify the impact of students' motivation to learn and behavioural engagement with course materials on subsequent training-related CT performance outcomes.



Students' perceived need for cognition (Cacioppo et al. 1984) was also examined, as research suggests that, in addition to motivation to learn, dispositional need for cognition is also significantly correlated with CT performance (Halpern 2006; Jensen 1998; King and Kitchener 2002; Toplak and Stanovich 2002). Thus, we hypothesised that CT performance would be positively correlated with both dispositional need for cognition and motivation towards learning at both pre-and-post-testing. We also examined whether or not any increases in need for cognition or motivation, from pre- to post-testing might account for gains in CT ability over and above the effects of training.

## Method

### Participants

Participants were first year psychology students, aged between 18 and 25 years, from the National University of Ireland, Galway. Two-hundred and forty-seven students (173 females, 74 males) expressed an interest in participating and attempted the online pre-tests. However, only 156 (108 females, 48 males) completed pre-testing; and only 74 (47 females, 27 males) completed post-testing. Non-completers reported not having enough time, principally as a result of having a heavy workload in other mandatory courses, as the primary reason for why they withdrew. There were no baseline differences (i.e. in either CT, need for cognition or motivation) between completers and non-completers. In return for their participation, students were awarded academic course credits. To ensure confidentiality, participants were identified by ID number only.

### Materials and measures

The materials made available during the CT course were the online lectures, exercises and feedback (see Table 2 and procedure below for more details). These materials are available upon request.

The Halpern Critical Thinking Assessment (HCTA; Halpern 2010) was administered at pre- and post-testing. The HCTA consists of 25 open-ended questions based on believable, everyday situations, followed by 25 specific questions that probe for the reasoning behind each answer. Questions on the HCTA represent five categories of CT applications: hypothesis testing (e.g. understanding the limits of correlational reasoning and how to know when causal claims cannot be made), verbal reasoning (e.g. recognising the use of pervasive or misleading language), argument analysis (e.g. recognising the structure of arguments, how to examine the credibility of a source and how to judge one's own arguments), judging likelihood and uncertainty (e.g. applying relevant principles of probability, how to avoid overconfidence in certain situations) and problem-solving (e.g. identifying the problem goal, generating and selecting solutions among alternatives). For an example of a question on the HCTA and how it is scored, see Fig. 2. Test reliability is robust; and ranges from 0.79 to 0.88 (Halpern 2010). The internal consistency of the scale in the current study was  $\alpha = .82$ .

The Need for Cognition Scale (Cacioppo et al. 1984) was administered at pre- and post-testing. The Need for Cognition (short form) consists of 18 items coded on a seven-point likert scale that assess one's willingness to explore and engage in relatively complex cognitive activities (e.g. "*I would prefer complex to simple problems*"; and "*I prefer to think about small, daily projects to long-term ones*"). The estimates of test reliability range

*Suppose that you are a first-year student in a dental school. You realize that your new friend, who is also a first-year student in dental school, is getting drunk on a regular basis several times a week. You do not see any signs of her drinking problem at school, but you are concerned because you will both begin seeing patients at the school's dental clinic within a month. She has not responded to your hints about her drinking problem. As far as you know, no one else knows about her excessive drinking.*

**Part A:** State the problem in two ways.

**Scoring:** There are two points possible for part A. Please answer the following question(s) in order to score the respondent's answers. Sum the scores from both questions.

Does the respondent's problem statement indicate that the new friend has a drinking problem and will be dealing with patients? Yes = 1 point; No = 0 points

Does the respondent's problem statement indicate that there are no signs that the drinking problem impairs performance? Yes = 1 point; No = 0 points

**Part B:** For each statement of the problem, provide two differed possible solutions.

**Scoring:** There are 2 sets of questions for part B. Two points are possible for each set of questions. Please answer the following question(s) in order to score the respondent's answers.

*Set 1:* Does the respondent suggest informing an authority figure about the problem? Yes = 2 points; No = 0 points

Does the respondent suggest that the friend should not deal with patients? Yes = 2 points; No = 0 points

*Set 2:* Does the respondent suggest showing the friend how the drinking problem could potentially impair her performance? Yes = 2 points; No = 0 points

Does the respondent suggest convincing the friend that she puts others in danger regardless of whether she knows it or not? Yes = 2 points; No 0 points

**Part C:** Given these facts, rate each of the following problem statements on a scale of 1 to 7 in which:

1 = extremely poor statement of the problem.

2 = very poor statement of the problem.

3 = poor statement of the problem.

4 = statement of the problem that is medium in quality.

5 = good statement of the problem.

6 = very good statement of the problem.

7 = excellent statement of the problem.

1. The friend may cause harm to patients because she is drunk.
2. You are the only one who knows she has a drinking problem.
3. Your friend's parents do not know she has a drinking problem.
4. You need to find a way to give your friend better hints about her drinking.
5. The friend may flunk out of school if she continues to get drunk so often.
6. The friend may hurt herself if she continues to get drunk so often.
7. You feel responsible for your friend's drinking problem.

**Scoring:** There are seven points possible in part C; one point is possible per question. If the respondent selected any number within the correct range they earn one point. If the respondent selected a number outside the correct range they do not earn a point.

Question 1: Correct range: 5-7; Question 2: Correct range: 2-5

Question 3: Correct range: 1-2; Question 4: Correct range: 1-4

Question 5: Correct range: 4-7; Question 6: Correct range: 4-7

Question 7: Correct range: 4-7

**Fig. 2** Question 21 on the HCTA (of the problem-solving sub-scale) with scoring protocol (Halpern 2010)



from 0.85 to 0.90 (Sherrard and Czaja 1999); and the internal consistency of the scale in the current study was  $\alpha=.91$ .

The Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich et al. 1991) was administered in order to match experimental and control groups on motivation at the pre-testing stage and to assess differences between groups at the post-testing stage, as it has been speculated that motivation influences the willingness to engage in metacognitive processes, such as CT (Ennis 1998; Dwyer et al. 2011; Garcia et al. 1992). The version of the MSLQ used in this research consisted of 43 items (e.g. *If I try hard enough, then I will understand the course material*; and *I work hard to do well in class even if I don't like what we are doing*), each of which is responded to using a seven-point likert scale (e.g. 1=strongly agree, 7=strongly disagree). Eight sub-scales of the MSLQ were used in this study (i.e. motivation towards: elaboration, critical thinking, effort regulation, metacognitive self-regulation, organisation, control of learning beliefs, and both intrinsic and extrinsic goal orientation). Internal consistency for sub-scales ranged from  $\alpha=0.65-0.88$ .

## Procedure

The study took place over a period of 8 weeks. Two groups took part in this study: those who participated in the e-learning CT course taught through AM (the AM group) and a control group (i.e. those who received no CT intervention). The AM group completed a six-week online CT course in which they viewed classes twice per week; completed two exercise sessions per week; and received detailed feedback for both exercises at the end of the week. Each class involved presenting the educational material to students through AMs. The exercises involved the manipulation of AMs and completion of relevant CT tasks using AMs.

Notably, a number of researchers in the field of CT argue that CT is a domain-general process, in that it can be instructed in any educational setting and applied to any academic subject (Abrami et al. 2008; Ennis 1998; Halpern 2003). For example, a recent meta-analysis conducted by Abrami et al. (2008) investigated the effects of different CT instruction methods, using Ennis' (1989) typology of four CT courses (i.e. general, infusion, immersion and mixed). In the general approach to CT training, actual CT skills and dispositions "are learning objectives, without specific subject matter content" (Abrami et al. 2008, p. 1105). The infusion of CT into a course requires specific subject matter content upon which CT skills are practiced. In the infusion approach, the objective of teaching CT within the course content is made explicit. In the immersion approach, like the infusion approach, specific course content upon which critical thinking skills are practiced is required. However, CT objectives in the immersed approach are not made explicit. Finally, in the mixed approach, critical thinking is taught independently of the specific subject matter content of the course. Results revealed that the typologies that made the CT element of instruction explicit to students (i.e. general, infusion and mixed) yielded the largest effects.

The results indicated that it is the nature in which CT is instructed (rather than the subject area in which CT is applied), that is the crucial element in aiding students develop CT skills. According to Ennis (1998) and Abrami and colleagues (2008), in order to optimally teach CT, the CT element of instruction must be made explicit to students (e.g. through a mixed or infused approach; Abrami et al. 2008; Ennis 1998); that is, making CT objectives and requirements clear to students plays an important role in the development of CT ability. Thus, given both the domain-generality of CT and that participants used in this research were 1st Year Arts students, it was decided that presenting topics subject to CT application should vary across academic domains within

the Arts programme (e.g. English, Philosophy, Psychology and Sociology) in order to: (1) avoid alienating any specific group or groups of students; and (2) maintain the interest of all students (regardless of academic field).

Students who participated in the course used the *Rationale*<sup>TM</sup> AM software made available to them for purposes of completing their exercises, and they were also encouraged to practice using the *Rationale*<sup>TM</sup> programme outside of the course environment. The conceptualisation of CT taught in our course was largely based on findings reported by Facione (1990) in the *Delphi Report*. The Delphi panel overwhelmingly agreed (i.e. 95 % consensus) that *analysis*, *evaluation* and *inference* were the core skills necessary for CT (Facione 1990; see Table 1 for the description of each skill provided by the Delphi Report). Notably, questions on the HCTA reflect the need for CT skills of analysis, evaluation and inference. Like the AM group, the control group attended their 1st year Arts lectures (e.g. History, English, Psychology, Philosophy and Sociology), but did not participate in the CT course in any manner (i.e. control group participants did not partake in any CT exercises or view any CT lectures, feedback or course materials).

In Week 1, the Need for Cognition Scale, the MSLQ, and the HCTA were administered prior to the commencement of the course. Participants were then randomly assigned to either the experimental or control groups. The course began in Week 2.

The e-learning classes were voice recorded and dubbed over a PowerPoint<sup>TM</sup> slideshow using CamTasia<sup>TM</sup> recording software. Classes lasted a maximum of 15 min each, as research has shown that didactically teaching students for longer than 15 min can substantially decrease attention to the source of instruction (Wankat 2002). In each class, students were taught to use CT skills via worked examples (in the form of AMs). Students were able to pause, rewind, and restart the class at anytime they wished. Immediately after each class, students were asked to complete a set of active learning AM exercises and email the completed exercises back to the primary investigator. Engagement in the course was measured according to the number of exercises emailed to the primary investigator. Again, the course outline and what was taught in each class is presented in Table 2. Feedback was provided to students at the end of each working week, that is, after they had completed and returned two set of exercises. Feedback focused on the structure of arguments provided by students; inferential relationships among propositions in their arguments; and the relevance and credibility of the propositions they used. Sample feedback (for exercises from Lecture 3.1) can be found in Appendix A. In Week 8, after the completion of the CT course, the HCTA, the MSLQ and the Need for Cognition Scale were administered as post-tests.

**Table 1** Core CT skills according to the Delphi Report (adapted from Facione 1990)

Skill	Description
Analysis	To identify the intended and actual inferential relationships among statements, questions, concepts, descriptions or other forms of representation intended to express beliefs, judgments, experiences, reasons, information, or opinions. Examining ideas: to determine the role various expressions play or are intended to play in the context of argument, reasoning or persuasion; to compare or contrast ideas, concepts, or statements; to identify issues or problems and determine their component parts, and also to identify the conceptual relationships of those parts to each other and to the whole. Detecting arguments given a set of statements or other forms of representation, to determine whether or not the set expresses, or is intended to express, a reason or reasons in support of or contesting some claim, opinion or point of view.

**Table 1** (continued)

Skill	Description
	<p>Analysing arguments: given the expression of a reason or reasons intended to support or contest some claim, opinion or point of view, to identify and differentiate: (a) the intended main conclusion, (b) the premises and reasons advanced in support of the main conclusion, (c) further premises and reasons advanced as backup or support for those premises and reasons intended as supporting the main conclusion, (d) additional unexpressed elements of that reasoning, such as intermediary conclusions, non-stated assumptions or presuppositions, (e) the overall structure of the argument or intended chain of reasoning, and (f) any items contained in the body of expressions being examined which are not intended to be taken as part of the reasoning being expressed or its intended background.</p>
Evaluation	<p>To assess the credibility of statements or other representations which are accounts or descriptions of a person's perception, experience, situation, judgment, belief, or opinion; and to assess the logical strength of the actual or intended inferential relationships among statements, descriptions, questions or other forms of representation.</p> <p>Assessing claims: to recognize the factors relevant to assessing the degree of credibility to ascribe to a source of information or opinion; to assess the contextual relevance of questions, information, principles, rules or procedural directions; to assess the acceptability, the level of confidence to place in the probability or truth of any given representation of an experience, situation, judgment, belief or opinion.</p> <p>Assessing arguments: to judge whether the assumed acceptability of the premises of an argument justify one's accepting as true (deductively certain), or very probably true (inductively justified), the expressed conclusion of that argument; to anticipate or to raise questions or objections, and to assess whether these point to significant weakness in the argument being evaluated; to determine whether an argument relies on false or doubtful assumptions or presuppositions and then to determine how crucially these affect its strength; to judge between reasonable and fallacious inferences; to judge the probative strength of an argument's premises and assumptions with a view toward determining the acceptability of the argument; to determine and judge the probative strength of an argument's intended or unintended consequences with a view toward judging the acceptability of the argument; to determine the extent to which possible additional information might strengthen or weaken an argument.</p>
Inference	<p>To identify and secure elements needed to draw reasonable conclusions; to form conjectures and hypotheses; to consider relevant information and to deduce the consequences flowing from data, statements, principles, evidence, judgments, beliefs, opinions, concepts, descriptions, questions, or other forms of representation.</p> <p>Querying evidence: in particular, to recognize premises which require support and to formulate a strategy for seeking and gathering information which might supply that support; in general, to judge that information relevant to deciding the acceptability, plausibility or relative merits of a given alternative, question, issue, theory, hypothesis, or statement is required, and to determine plausible investigatory strategies for acquiring that information.</p> <p>Conjecturing alternatives: to formulate multiple alternatives for resolving a problem, to postulate a series of suppositions regarding a question, to project alternative hypotheses regarding an event, to develop a variety of different plans to achieve some goal; to draw out presuppositions and project the range of possible consequences of decisions, positions, policies, theories, or beliefs.</p> <p>Drawing conclusions: to apply appropriate modes of inference in determining what position, opinion or point of view one should take on a given matter or issue; given a set of statements, descriptions, questions or other forms of representation, to deduce, with the proper level of logical strength, their inferential relationships and the consequences or the presuppositions which they support, warrant, imply or entail; to employ successfully various sub-species of reasoning, as for example to reason analogically, arithmetically, dialectically, scientifically, etc.; to determine which of several possible conclusions is most strongly warranted or supported by the evidence at hand, or which should be rejected or regarded as less plausible by the information given.</p>

**Table 2** e-Learning CT course outline

Class no.	Title	What was taught
1	Pre-testing	• Students completed the HCTA, MSLQ and Need for Cognition post-test
2	Classes 1 and 2: “Introduction to Critical Thinking”	<ol style="list-style-type: none"> <li>1. We think in order to decide what to do and what to believe.</li> <li>2. We ultimately decide what to believe by adding supports or rebuttals to our own arguments (i.e. questioning our own beliefs).</li> <li>3. Arguments are hierarchical structures. We can continue to add more levels if we like.</li> </ol>
3	Classes 3 and 4: “Unpacking (analysing and evaluating) a persons’ belief”	<ol style="list-style-type: none"> <li>1. In order to analyse an argument, we must extract the structure of the argument from dialogue or prose.</li> <li>2. Identifying types (sources) of arguments and considering the strength of each type is another form of analysis.</li> <li>3. The evaluation of the overall strengths and weaknesses of an argument can be completed after adequate analysis.</li> </ol>
4	Classes 5 and 6: “Analysis & Evaluation”	<ol style="list-style-type: none"> <li>1. Evaluation includes the recognition of imbalances, omissions and bias within an argument.</li> <li>2. Evaluative techniques can aid recall.</li> <li>3. Examining whether or not the arguments used are <i>relevant or logically connected</i> to the central claim is also an important factor in evaluation.</li> </ol>
5	Classes 7 and 8: “Evaluation”	<p>We must evaluate:</p> <ol style="list-style-type: none"> <li>1. Types (sources) of arguments based on credibility</li> <li>2. The relevance of propositions to the central claim or intermediate conclusions within the argument</li> <li>3. The logical strength of an argument structure</li> <li>4. The balance of evidence within an argument structure</li> </ol>
6	Classes 9 and 10: “Inference”	<ol style="list-style-type: none"> <li>1. Evaluation and inference are intimately related.</li> <li>2. Inference differs from evaluation in that the process of inference involves <i>generating</i> a conclusion from previously evaluated propositions.</li> <li>3. In larger informal argument structures, intermediate conclusions must be inferred prior to the inference of a central claim.</li> </ol>
7	Classes 11 and 12: “Reflective Judgment”	<ol style="list-style-type: none"> <li>1. Reflective judgment is our ability to reflect upon what we know and the knowledge the world presents us; and our ability to think critically and reflectively in this context.</li> <li>2. One’s understanding of the nature, limits and certainty of knowing and how this can affect our judgment.</li> <li>3. Recognition that some problems cannot be solved with absolute certainty (i.e. ill-structured problems).</li> <li>4. The importance of structure and complexity in reflective judgment.</li> </ol>
8	Post-testing	• Students completed the HCTA, MSLQ and Need for Cognition post-tests

## Results

### Analysis

With respect to group differences in CT, need for cognition, and motivation, a series of 2 (time: pre-and-post-testing) x 2 (condition: AM group and control group) Mixed ANOVAs

were conducted to examine the effects of both time and CT training condition on need for cognition and motivation. A series of independent sample t-tests was used to compare the intervention and control groups on CT gain from pre-to-post-testing. A series of matched sample t-tests examined the effects of the CT intervention on CT and CT sub-skills ability from pre-to-post-testing. With respect to high-engagement versus low engagement differences in the AM group, a series of independent sample t-tests compared the gains in CT performance, from pre- to post-testing, of those who had a high level of engagement with the CT intervention with those who had a low level of engagement. Furthermore, Pearson correlations among CT performance, need for cognition and motivation sub-scales (i.e. at both pre- and post-testing) were also conducted.

Means and standard deviations for performance scores of the AM, control, high engagement and low engagement groups on overall CT, all CT sub-scales, motivation and need for cognition are presented in Table 3.

#### Group differences in critical thinking, need for cognition, and motivation

A series of 2 (time: pre-and-post-testing) x 2 (condition: AM group and control group) Mixed ANOVAs were conducted to examine the effects of both time and CT training condition on need for cognition and motivation. Results revealed that there was no effect of condition or time on need for cognition, and no condition x time interaction effect. Similarly, there was no effect of condition on motivation, and no condition x time interaction effect. However, there was a main effect of time on motivation,  $F(1, 65)=8.63$ ,  $p=.005$ , partial  $\eta^2=.12$ , with total motivation scores decreasing from pre-test ( $M=175.15$ ;  $SD=27.20$ ) to post-test ( $M=168.73$ ;  $SD=31.00$ ) in the sample as a whole. Post-hoc analyses also revealed that there was a significant decrease from pre-to-post-testing in two of the eight subscales: meta-cognitive self-regulation ( $t=2.13$ ,  $df=66$ ,  $p=.039$ , two tailed,  $d=.21$ ) and effort regulation ( $t=4.21$ ,  $df=66$ ,  $p<.001$ , two tailed,  $d=.34$ ).

A series of independent sample t-tests was used to compare the intervention and control groups on CT gain from pre-to-post-testing. Results revealed that students in the AM group showed significantly higher gain than controls for overall CT ability ( $t=-2.43$ ,  $df=68$ ,  $p<.018$ , two tailed,  $d=.60$ ) and for the sub-scale argument analysis ( $t=-2.29$ ,  $df=68$ ,  $p=.025$ , two tailed,  $d=.54$ ). There were no differences in gain between the groups on any other CT sub-scale.

A series of matched sample t-tests examined the effects of the CT intervention on CT and CT sub-skills ability from pre-to-post-testing. Results revealed that students in the AM group scored significantly higher on post-testing compared with pre-testing on overall CT ability ( $t=-6.65$ ,  $df=42$ ,  $p<.001$ , two tailed,  $d=.81$ ) and all CT sub-scales: hypothesis testing ( $t=-3.89$ ,  $df=41$ ,  $p<.001$ , two tailed,  $d=.55$ ), verbal reasoning ( $t=-2.97$ ,  $df=41$ ,  $p=.005$ , two tailed,  $d=.49$ ), argument analysis ( $t=-2.14$ ,  $df=42$ ,  $p=.038$ , two tailed,  $d=.40$ ), likelihood/uncertainty ( $t=-4.64$ ,  $df=41$ ,  $p<.001$ , two tailed,  $d=.67$ ) and problem-solving ( $t=-4.47$ ,  $df=41$ ,  $p<.001$ , two tailed,  $d=.64$ ). Results further revealed that students in the control group scored significantly higher on the post-test than on the pre-test on overall CT ability ( $t=-3.01$ ,  $df=30$ ,  $p=.005$ , two tailed,  $d=.55$ ) and also for the sub-scale of problem-solving ( $t=-3.77$ ,  $df=27$ ,  $p=.001$ , two tailed,  $d=.65$ ).

#### High-engagement versus low-engagement differences in the AM group

A series of independent sample t-tests compared the gains in CT performance, from pre- to post-testing, of those who had a high level of engagement (exercises

**Table 3** Means and standard deviations for CT performance by condition

	N	Pre-test		Post-test	
		M	SD	M	SD
Overall critical thinking					
AM (High engagement)	19	96.26	10.98	110.21	15.16
AM (Low engagement)	23	96.54	12.83	106.88	14.23
AM (total)	42	98.00	11.01	109.12	13.83
Control	28	94.50	12.54	99.79	13.66
Hypothesis testing					
AM (High engagement)	19	21.74	3.83	24.58	4.32
AM (Low engagement)	23	22.70	4.29	24.65	4.65
AM (total)	43	22.26	4.07	24.44	4.55
Control	31	21.10	4.53	21.90	5.96
Verbal reasoning					
AM (High engagement)	19	6.74	1.52	7.74	2.58
AM (Low engagement)	23	6.13	2.30	7.43	2.86
AM (total)	43	6.40	1.99	7.56	2.68
Control	31	5.71	1.94	6.03	2.35
Argument analysis					
AM (High engagement)	19	22.26	4.28	25.16	4.91
AM (Low engagement)	24	23.00	3.50	23.63	3.91
AM (total)	43	22.67	3.83	24.30	4.40
Control	31	22.00	5.01	21.00	4.58
Likelihood/uncertainty					
AM (High engagement)	19	9.53	3.70	11.53	3.79
AM (Low engagement)	23	10.52	3.51	13.35	3.55
AM (total)	43	10.07	3.59	12.44	3.73
Control	31	9.07	3.44	10.29	4.02
Problem-solving					
AM (High engagement)	19	36.00	3.71	41.21	5.71
AM (Low engagement)	24	37.00	4.37	38.74	5.26
AM (total)	43	36.42	4.10	39.36	3.70
Control	31	36.21	4.69	39.61	5.75
Need for cognition					
AM (High engagement)	19	66.16	14.69	68.00	16.53
AM (Low engagement)	24	67.67	15.29	66.67	16.73
AM (total)	43	67.00	14.87	67.26	16.45
Control	26	65.85	19.10	65.08	15.65
Motivation					
AM (High engagement)	19	176.16	30.75	173.16	27.29
AM (Low engagement)	24	166.50	26.22	157.73	37.66
AM (total)	41	170.98	28.46	164.88	33.77
Control	26	181.73	24.15	174.81	25.49



completed, range: 12–24;  $M=20.67$ ) with the CT intervention with those who had a low level of engagement (exercises completed, range: 0–11;  $M=6.82$ ). Results revealed that students in the high engagement group exhibited a significantly higher gain in CT performance from pre- to post-testing when compared with the low engagement group on the CT sub-scale of problem-solving ( $t=-2.95$ ,  $df=40$ ,  $p=.005$ , two tailed,  $d=.91$ ). There were no other differences on CT performance, need for cognition or motivation observed between the two engagement groups.

### Correlations

There was a significant correlation between need for cognition and motivation at pre-testing ( $r=.52$ ,  $p<.001$ ) and at post testing ( $r=.60$ ,  $p<.001$ ), but neither need for cognition nor motivation were correlated with CT performance at pre-testing. There was a significant correlation between CT performance and both need for cognition ( $r=.47$ ,  $p<.001$ ) and motivation ( $r=.28$ ,  $p=.017$ ) at post-testing. The full set of correlations among CT, need for cognition and motivation sub-scales at pre-testing and post-testing are presented in Table 4. Results from a regression analysis revealed that change in need for cognition ( $\beta=-.09$   $p=.510$ ) and motivation ( $\beta=.08$   $p=.541$ ) from pre-to-post-testing did not account for any variance (adjusted  $r^2=.02$ ) in CT gains over and above the effect of experimental condition,  $F(2, 62)=.71$ ,  $p=.749$ .

## Discussion

### Interpretation of results

The current study set out to examine, first, whether AM training delivered using an e-learning CT course would significantly enhance CT performance in comparison with a control condition; second, we tested the hypothesis that students who engaged more with the course would perform significantly better on CT performance than those who did not engage as much; and third, we examined the claim that CT performance would be positively correlated with both dispositional need for cognition and motivation at both testing times.

The results of the current study revealed that students in the AM group scored significantly higher on post-testing than on pre-testing on measures of overall CT ability and on all CT sub-scales. Results also revealed that those in the control condition improved from pre-to-post-testing on overall CT ability and the problem-solving sub-scale of the HCTA. It is possible that overall CT and problem-solving performance might improve over time due to maturation. Alternatively, it could be that improvements in the control group (and likewise the AM group) were as a result of practice effects (i.e. repeat administration of the HCTA). However, positive effects of the AM training course were observed over-and-above any possible maturation or practice effects. Specifically, results revealed that those in the AM group showed a significantly larger gain from pre- to post-testing than those in the control group on overall CT ability and the CT sub-scale of argument analysis. Given that there were no significant differences between the control and AM groups at the pre-testing stage on either overall CT or on any CT sub-scales, these findings suggest that the two groups were adequately matched on CT ability prior to the intervention and that participation in an e-learning CT course taught through AM significantly enhances CT performance.

We did not find strong evidence in favour of our second hypothesis. Specifically, there was no difference between those who engaged more (i.e. those who completed 12–24 CT

**Table 4** Correlations among CT performance, need for cognition and motivation sub-scales at pre-testing (below diagonal) and post-testing (above diagonal)

	CT	NFC	IGO	CoLB	EGO	Org.	MSR	EffReg	MotCT	Elab
CT performance (CT)	-	$r=.47$ $p<.001$	$r=.28$ $p=.016$	$r=.24$ $p=.045$	$r=-.07$ $p=.556$	$r=.02$ $p=.812$	$r=.24$ $p=.039$	$r=.15$ $p=.201$	$r=.36$ $p=.002$	$r=.40$ $p<.001$
Need for cognition (NFC)	$r=.17$	-	$r=.66$ $p<.001$	$r=.44$ $p<.001$	$r=.20$ $p=.094$	$r=.29$ $p=.013$	$r=.46$ $p<.001$	$r=.47$ $p<.001$	$r=.49$ $p<.001$	$r=.57$ $p<.001$
Intrinsic Goal Orientation (IGO)	$r=.01$	$r=.50$	-	$r=.52$ $p<.001$	$r=.20$ $p=.098$	$r=.48$ $p<.001$	$r=.57$ $p<.001$	$r=.46$ $p<.001$	$r=.67$ $p<.001$	$r=.60$ $p<.001$
Control of learning beliefs (CoLB)	$r=.12$	$r=.40$	$r=.56$	-	$r=.34$ $p=.004$	$r=.29$ $p=.015$	$r=.40$ $p<.001$	$r=.40$ $p<.001$	$r=.38$ $p=.001$	$r=.48$ $p<.001$
Extrinsic goal orientation (EGO)	$r=.03$	$r=.21$	$r=.30$	$r=.37$	-	$r=.28$ $p=.016$	$r=.16$ $p=.193$	$r=.15$ $p=.208$	$r=.13$ $p=.262$	$r=.12$ $p=.302$
Organisation (Org)	$r=.801$	$p=.091$	$p=.013$	$p=.002$	$r=.16$	-	$r=.67$ $p<.001$	$r=.56$ $p<.001$	$r=.53$ $p<.001$	$r=.68$ $p<.001$
Metacognitive self-regulation (MSR)	$r=-.14$	$r=.30$	$r=.38$	$r=.18$	$r=.177$	$r=.65$	$r=.65$ $p<.001$	$r=.73$ $p<.001$	$r=.72$ $p<.001$	$r=.71$ $p<.001$
Effort regulation (EffReg)	$p=.237$	$p=.013$	$p=.001$	$p=.133$	$p=.094$	$p=.50$	$r=.63$ $p<.001$	$r=.73$ $p<.001$	$r=.59$ $p<.001$	$r=.64$ $p<.001$
Motivation towards CT (MotCT)	$r=.03$	$r=.36$	$r=.50$	$r=.30$	$r=.20$	$r=.50$	$r=.63$ $p<.001$	-	$r=.59$ $p<.001$	$r=.64$ $p<.001$
Elaboration (Elab)	$p=.828$	$p=.003$	$p<.001$	$p=.012$	$p=.094$	$r=.54$	$r=.62$ $p<.001$	$r=.33$ $p=.006$	$r=.59$ $p=.001$	$r=.67$ $p<.001$
	$r=-.02$	$r=.33$	$r=.37$	$r=.24$	$r=.35$	$r=.54$	$r=.62$ $p<.001$	$r=.33$ $p=.006$	-	$r=.73$ $p<.001$
	$p=.866$	$p=.006$	$p=.002$	$p=.046$	$p=.003$	$r=.54$	$r=.62$ $p<.001$	$r=.33$ $p=.006$	$r=.59$ $p<.001$	$r=.67$ $p<.001$
	$r=.04$	$r=.54$	$r=.52$	$r=.29$	$r=.12$	$r=.67$	$r=.71$ $p<.001$	$r=.51$ $p<.001$	$r=.73$ $p<.001$	-
	$p=.746$	$p<.001$	$p<.001$	$p=.017$	$p=.317$	$r=.198$	$r=.71$ $p<.001$	$r=.51$ $p<.001$	$r=.73$ $p<.001$	-
	$r=.14$	$r=.39$	$r=.48$	$r=.32$	$r=.16$	$r=.198$	$r=.71$ $p<.001$	$r=.51$ $p<.001$	$r=.73$ $p<.001$	-
	$p=.265$	$p=.001$	$p<.001$	$p=.007$	$p=.198$	$r=.198$	$r=.71$ $p<.001$	$r=.51$ $p<.001$	$r=.73$ $p<.001$	-

exercises) or less (i.e. those who completed 0–11 CT exercises) on overall CT ability. However, those in the high-engagement group did show a greater gain from pre-test to post-testing in problem-solving ability than those in the low-engagement group. According to the HCTA manual (Halpern 2010, p. 7), “problem-solving involves the use of multiple problem statements to define the problem and identify possible goals, the generation and selection of alternatives, and the use of explicit criteria to judge among alternatives.” Notably, problem-solving, as defined by Halpern (2010), is akin to the CT sub-skill of inference as defined in the *Delphi Report* (Facione 1990; again, see Table 1).

Although we found an effect of engagement on problem-solving in the AM group, this effect did not transfer to the overall CT ability. These results do not conform with the pattern of results reported in van Gelder et al. (2004), where it was found that level of AM practice in a computer-supported learning environment positively correlated with CT ability. However, one difference between our study and the study of van Gelder and colleagues is that the students in their training course did not view pre-recorded AM training lectures prior to practicing AM exercises; rather they logged into a practice portal and spent time working independently on AM projects. One possible explanation for the results in our study (i.e. an overall positive effect of AM training condition but a weaker effect of exercise engagement) may be due to the high quality of the lectures and/or the feedback (provided to all students in the AM group, regardless of whether or not they completed more or less exercises), which may have provided students with sufficient engagement with AM to improve CT ability. Also, given that the average level of engagement among all students who participated in the course was 11.67 (out of 24), with only 20.93 % of sample engaging with none of the exercises, it is important to note that the majority of students engaged beyond simply viewing online lectures, including the majority of students in the low-engagement group. It is also important to note that findings regarding level of engagement should be interpreted with caution given that engagement was not an experimentally manipulated variable.

Upon further analysis, the CT performance of those in the high engagement group was positively correlated with motivation ( $r=.38$ ;  $p=.037$ ), as well as with need for cognition ( $r=.49$ ;  $p=.005$ ). In addition, though it was found that the CT performance of those in the low engagement group was positively correlated with their need for cognition ( $r=.38$ ;  $p=.022$ ), CT performance was not correlated with motivation. These findings indicate that motivation was a key feature of performance in the high engagement group, with motivation and performance being more closely coupled in the high engagement group but not in the low engagement group.

We also explored the relationship between need for cognition, learning motivation, and CT ability in the sample as a whole. We observed that neither a significant relationship between CT performance and need for cognition existed, nor a significant relationship between CT performance and learning motivation existed, at pre-test. However, on post-test both need for cognition and learning motivation were significantly correlated with CT performance, suggesting that, consistent with previous research and theory (e.g. Ennis 1998; Halpern 2003, 2006; Ku and Ho 2010a), there is some inter-dependency between a positive or motivated disposition toward thinking and learning and CT ability.

While we observed a correlation between overall motivation to learn and CT ability, this positive interdependence was accounted for by five of the eight subscales: intrinsic goal orientation, control of learning beliefs, metacognitive self-regulation, motivation towards critical thinking, and motivation towards elaboration. Motivation towards elaboration (i.e. motivation to elaborate on information via paraphrasing, summarisation and/or creating analogies to build connections between different items of information; Pintrich et al. 1991) was the motivation sub-scale with the highest correlation with CT performance at

post-test, possibly due to good critical thinkers choosing to conduct a deeper analysis of the structure of arguments. Motivation towards CT (i.e. motivation to apply knowledge to new situations in order to make evaluations, solve problems and/or reach decisions) was the motivation sub-scale with the second highest correlation with CT performance, which is perhaps unsurprising, given that the HCTA requires the application of knowledge to problem situations in order to make evaluations and reach a decision in relation to key probe questions. However, results also indicated that motivation towards elaboration had a higher correlation with CT ability in the control group ( $r=.60$ ) than with the AM group ( $r=.35$ ), as did motivation towards CT (control:  $r=.37$ ; AM:  $r=.35$ ); which suggests that CT training was not the critical factor binding motivation and CT ability over time. Outside of the pre-screening experience itself, or a post-screening reflection period, the novel learning experience of the first year at university was the only other significant factor that may have caused the increased coupling of motivation and ability over time, and this is perhaps unsurprising as most first year (i.e. freshman) courses challenge students to think in new and different ways, and critically, about information that is being presented to them in lectures.

Although the relationship between need for cognition, learning motivation, and CT ability changed from pre-test to post-test, the results of the current study revealed that there was no effect of AM training on average levels of need for cognition or learning motivation. This finding suggests that differences between the AM and control groups on CT performance were not caused by changes in students' dispositional need for cognition or motivation to learn. The results of regression analysis further clarified that change in need for cognition and motivation from pre-to-post-testing did not account for any variance (adjusted  $r^2=.02$ ) in CT gains over and above the effect of experimental condition.

Results further indicated that though students' need for cognition did not change over time in either the experimental or control groups, their motivation to learn significantly decreased over time. Upon closer analysis it was found that this global reduction in motivation was accounted for by a significant reduction in two of the eight motivation sub-scales: effort regulation and metacognitive self-regulation. Notably, effort regulation was not correlated with CT ability. In addition, though metacognitive self-regulation was correlated with CT ability, the correlation was moderate at best (i.e.  $r=.24$ ). Given that the experimental and control groups did not differ in this regard, the decrease in motivations may not be indicative of the CT training course and may have been a result of some factor outside of the course itself. This may simply be an effect of time, with a general decrease in motivation observed in first year Arts from early to late in the first semester. Specifically, it may be that in general, students' workload demands had a negative influence on effort regulation over time; or perhaps the novelty of being in college 'wore off' and students began to lose interest in maintaining their initial levels of effort.

#### Limitations of the current research

Though this study revealed that CT performance can be significantly enhanced by participation in an e-learning CT course taught through AM, there are some limitations that must be considered. One limitation of the current study was the sample size available for analysis after completion of the intervention. With a prospective pool of approximately 1,200 first year undergraduate students, 720 of whom were psychology students, only 156 students participated in the pre-testing session and only 74 completed post-testing after the intervention period. However, it is also worth noting that there were no differences between

completer and non-completers on need for cognition or on motivation scores at the pre-testing stage. These findings indicate that a lack of motivation or dispositional need for cognition was not accounting for attrition and that attrition may be more dependent on other personal factors (e.g. personal reasons, not having enough time to take part, or preferring to use any extra time between mandatory lectures for study). We speculate that the relatively small sample size (the result of attrition) may have impacted on the power of our statistical analysis and may also have accounted for the null findings associated with level of engagement (i.e. the sample size of the engagement analysis in the AM group was only 43). In addition, the attrition of students from pre- to post-testing resulted in differences between the groups with regard to the post-intervention sample sizes available for analysis (i.e. AM group:  $N=43$ , Control group:  $N=31$ ) and this in turn might have been at least partially responsible for the differences in CT abilities at post-testing.

In order to overcome problems of attrition, future research might aim to implement and evaluate CT interventions in the context of a mandatory course, as opposed to a voluntary course (as in this research). Although psychology students who participated in this study were promised credits towards their overall 1st Year Psychology mark, and all study participants were offered a certificate of completion and the possibility of winning a cash prize, it seems that this was not enough to keep all participants involved in the study. By making our CT intervention mandatory, attrition would have been significantly reduced and motivation (i.e. metacognitive self-regulation and effort regulation) may have increased rather than decreased over time.

Though the AM group was compared with a control group, another limitation of this study is that it was not compared with another CT training condition. While including a control group for comparison purposes is important for all CT intervention studies, given the hypothesized value of AM training as a means of promoting the development of CT skills, it is important that future research include other active control conditions that involve training of CT skills using more traditional means, or alternate conditions where AM practice, type of feedback, or course delivery strategy is manipulated. For example, although Alvarez-Ortiz's (2007) meta-analysis found that courses where there was "lots of argument mapping practice" (LAMP) produced a significant gain in students' critical thinking performance, with an effect size of .78 SD, CI [.67, .89], students who participated in CT courses that used at least some argument mapping within the course achieved gains in CT ability with an effect size of .68 SD, CI [.51, .86]. Thus, while the amount of AM practice may be a significant variable worth manipulating in future intervention comparison studies, it appears that researchers will need to think carefully about how to maximize the benefits of AM practice.

One final limitation of the current study that should be considered was the relatively low internal consistency of some scales on the MSLQ. This may have potentially influenced some observed correlations and/or reliability of estimates. Though some caution may be necessary in interpreting these findings, nevertheless, the correlations reported remain consistent with previous research on the relationship between motivation and CT (e.g. Garcia et al. 1992; Valenzuela et al. 2011) and can be viewed as providing further support for such claims.

#### Future research

Future research on CT interventions could also move beyond measuring CT performance according to quantitative assessment, to include qualitative analyses of how students come to answer CT questions/problems. For example, in research by Ku and Ho (2010b), it was found that when asked to 'talk aloud' when critically thinking about each question on the

HCTA, students who were proficient at CT engaged in more metacognitive activities and processes, including self-regulatory planning and evaluation skills. Future research could potentially examine the effect of argument mapping on the structure of metacognitive processing during CT ‘think aloud’ protocols and how these processes influence CT performance. Developing Ku and Ho’s line of enquiry, this deeper qualitative analysis of the benefits of AM training may also shed light on the relationship between metacognitive processes, such as self-regulatory planning, and the increase (or decrease) in disposition toward critical thinking and the coupling (or decoupling) of disposition and ability over time.

In addition, given that research suggests that feedback provided during AM training can enhance CT ability (Butchart et al. 2009; van Gelder 2003), future research could also examine the effects of specific types of AM training feedback on different aspects of critical thinking ability. For example, future research might examine the effects of feedback focused specifically on students’ ability to analyse the credibility of propositions in comparison with feedback focused specifically on students’ ability to evaluate the relevance of propositions, or the inferential relationships and logical strength of proposition structures. Moreover, given that both motivation towards learning and need for cognition were significantly correlated with CT performance at post-testing, future research should also take care to control for both of these variables as differences in CT performance that emerge as a result of interventions, regardless of design, may potentially be accounted for by differences in either motivation or need or cognition.

Furthermore, given (1) the potentially beneficial effects of AM in e-learning environments (as observed in the current research), (2) that AM is used to visually represent the structure of arguments and allows for their manipulation (van Gelder and Rizzo 2001; van Gelder 2003) and (3) that argumentation is a social activity (van Eemeren et al. 1996), it seems reasonable to further speculate that the ability of computer-supported AM to enhance CT may be optimised in collaborative learning settings. It has been argued by Paul (1987, 1993) that dialogue, a fundamental component of collaborative learning, is necessary for good CT. In the context of CT and computer-based AM, dialogue is advantageous because it provides thinkers with an opportunity to explain and question their own beliefs and arguments in light of the thinking and opinions of others involved in the dialogue. In this way, the thinkers involved in the dialogue are actively engaged in collaborative learning. Past research indicates that the use of mapping strategies (e.g. argument mapping) in computer-supported collaborative learning environments can facilitate: (1) higher grades on academic course assessments; (2) reasoned discussion among students; and (3) aid in focusing students to transfer these dialogic skills to curriculum-based learning (Engelmann et al. 2010; Engelmann and Hesse 2010; Hwang et al. 2011; Johnson et al. 2000; Ma 2009; Wegerif and Dawes 2004). Research has shown that reasoning and argumentation skills increase if computer-supported collaborative learning environments is used, given that it aids students in making their thoughts and solution strategies clear (Kuhn et al. 2008; Wegerif 2002; Wegerif and Dawes 2004). These recommendations for future research regarding AM are further supported by recent research which suggests that collaborative learning through mapping strategies, similar to AM, can enhance learning performance (Engelmann and Hesse 2010; Engelmann et al. 2010; Hwang et al. 2011; Roth and Roychoudhury 1994).

A more global perspective on the findings from this research suggests that AM can potentially supplement traditional methods of presenting arguments that are the focus of CT. For example, based on the findings of the current research and previous research in our laboratory (Dwyer et al. 2010), it appears that AM can be successfully used: (1) to support didactic instruction or to potentially replace text-based learning strategies in certain situations; (2) as a study guide provided by the teacher to be used by the student, (3) as a partially



completed study guide provided by the teacher to be completed by students when reading text, and/or (4) as a means of providing students with a method of constructing arguments from scratch using specific, class-based material as the basis for AM construction work.

Specifically, in didactic, instructional settings, instead of presenting students with slideshows filled with bullet points of information that they will need to recall in the future, it may prove more advantageous to place AMs within the slides as a means of presenting both the target information and the structure of the reasoning behind it. In this context, AMs may provide students with the opportunity to gain deeper insight and greater understanding of the subject being taught, through assimilating the propositions, drawing the necessary connections among those propositions, and assessing the relevance, credibility and the logical strength of those propositions and their interconnectivity within an AM. Thus, as a result of potentially greater understanding and deeper insight, students may be better able to better analyse and evaluate the class materials. That is, AMs may provide students with a visual scaffold of the information expected to be learned; and may also aid in their ability to critically analyse and evaluate the target information for purposes of creating greater understanding.

Presenting information in this hierarchically organised manner may also allow students to more readily question the importance of propositions and their relationships within class materials, given that the structure of the information is made explicit; and may possibly motivate students to seek further justification from sources apart from class-based materials. That is, if an argument is explicitly laid out for students in class via an AM, it may facilitate their ability to see the logical flow of the argument more easily, given that they are spared the need to simultaneously assimilate the argument and take notes. Thus, the use of AMs in the classroom may promote student engagement in the classroom.

In addition to the benefits of AM in didactic settings, the ability to actively map arguments could potentially aid students to organise their notes outside of the classroom and more easily assimilate important information from additional readings. This in turn would allow them to actively learn information through their own investigation of the given subject area. Furthermore, findings from the current research indicate that AM 'know-how' provides students with the opportunity to actively learn, in that students are provided a means of structuring propositions into arguments, gathered from both classroom-based and extracurricular investigations, for purposes of analysing and evaluating the materials and inferring their own conclusions; thus providing them with the opportunity to actively gain a deeper understanding of the subject area.

In conclusion, consistent with reports which highlight the value of using e-learning to facilitate the development of metacognitive processes and active learning (Huffaker and Calvert 2003), the results of the current study suggests that CT skills can be enhanced by participating in an AM training course delivered in an e-learning environment. However, future research is necessary to further examine the conditions that most positively affect CT and dispositions towards thinking.

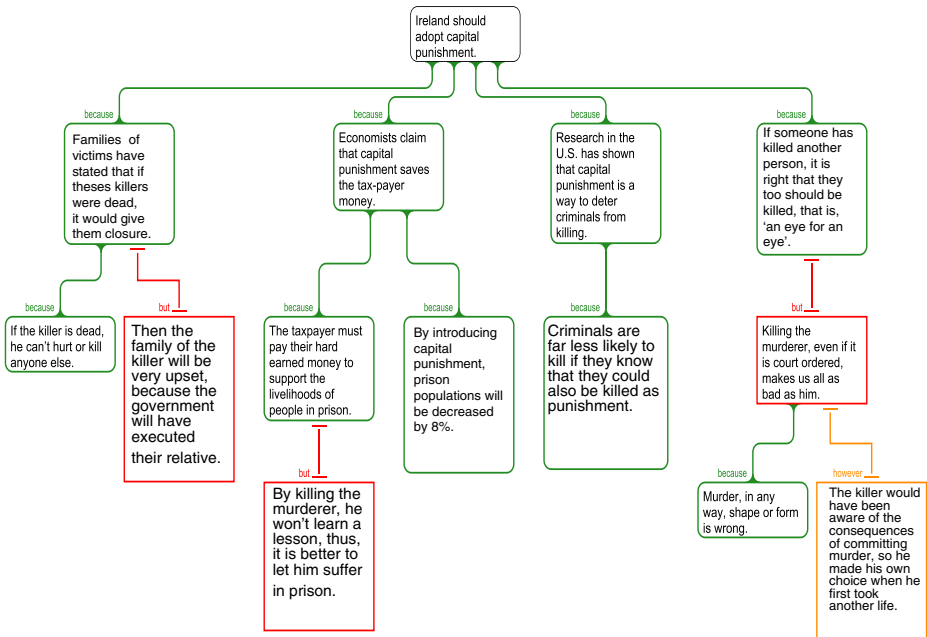
## Appendix A: sample feedback from lecture 3.1

Thank you very much to all of you who did the exercises. Below you will find some feedback on exercises from Lecture 3.1.

Exercise: "Ireland should adopt Capital Punishment"

Below please find the argument we have extracted from the text you were asked to analyse. Please compare and contrast your argument map with this one. Also, please take

note of where you may have differed in your placement of some propositions and why you made the analysis decisions you made.



You were also asked to answer a few questions based on this argument map.

Question 1 asked:

Does the author sufficiently support their claims? Are the author’s claims relevant? Does the author attempt to refute their own arguments (i.e., disconfirm their belief)?

Some of you answered that:

The author did not sufficiently support his/her claims because most of the propositions used were based on personal opinion (i.e. there was an insufficient amount of evidence to suggest that Ireland should adopt capital punishment). The author did attempt to refute their own claims, however, did so poorly in that he/she only used personal opinion or ‘common belief’ statements.

The author sufficiently supported their claims as they were sufficiently backed up by other supports. These claims are all relevant to the argument, specifically the central claim. The author does not attempt to disconfirm his beliefs because he sticks to his guns that capital punishment should be adopted.

The truth of the matter is that the author *did not sufficiently* support his/her claims. Of the 8 reasons he provided, only 3 were based on either expert opinion, statistics from research or research data.

All the arguments made were relevant to the central claim.

The author *did attempt* to refute his/her claims (i.e. disconfirm their own belief), as on 3 occasions, some form of objection to the reasoning was presented. However, the objections used were not examples of high quality evidence.

Question 2 asked:

Are there other arguments you would include?

Some of you answered that:

Some argument should be made in terms of when the death penalty should/would be used, such as in cases of mental problems or a conviction of manslaughter.

Some argument should consider the nature of the crime, such as how the murder took place –details should be considered.

Everyone has a right to life, even murderers.

Law abiding citizens might grow to fear the government as they would now have more control over you.

Please think about these ideas and claims and also think about how you could possibly integrate them into the argument map. In addition, think about how you might support or object to these new propositions.

Question 3 asked:

Does any proposition or any set of propositions suggest to you that the author is biased in any way?

Some of you answered:

The author is biased because he/she presents more reasons for why we should adopt capital punishment than for not adopting capital punishment.

The author was not biased because though he/she did present more reasons in favour of capital punishment, they were mostly based on personal opinion and were adequately objected to.

The author does certainly appear to be biased. However, some of you argued that it is because the author stated that ‘Ireland should adopt capital punishment’, thus making it a biased argument from the outset. This is not true, because the author may have made the same claim and then simply presented 5 objections at level 1 in the argument structure (as opposed to 4 supports). Remember, there is more to determining bias than simply assimilating what the central claim is; what is more important is how the author attempts to justify or refute this claim. The reason why this argument is biased is because the author only presents some arguably credible evidence (in 3 cases) to support the claim. In other cases where the author makes a claim and objects to it, both the reasons and objections are based on personal or common belief. This is done to disguise the author’s bias. In the cases where the author presents credible evidence, there are no objections.

## References

- Abrami, P. C., Bernard, R. M., Borokhovski, E., Wade, A., Surkes, M. A., Tamim, R., & Zhang, D. (2008). Instructional interventions affecting critical thinking skills and dispositions: a stage 1 meta-analysis. *Review of Educational Research*, 78(4), 1102–1134.
- Alvarez-Ortiz, C. (2007). *Does philosophy improve critical thinking skills?* Unpublished thesis. The University of Melbourne.
- Association of American Colleges & Universities (2005). *Liberal education outcomes: A preliminary report on student achievement in college*. Washington, DC.
- Australian Council for Educational Research (2002). *Graduate skills assessment*. Commonwealth of Australia.
- Boekaerts, M., & Simons, P. R. J. (1993). *Learning and instruction: Psychology of the pupil and the learning process*. Assen: Dekker & van de Vegt.
- Brown, A. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In F. Reiner & R. Kluwe (Eds.), *Metacognition, motivation, and understanding* (pp. 65–116). Hillsdale: Erlbaum.
- Butchart, S., Bigelow, J., Oppy, G., Korb, K., & Gold, I. (2009). Improving critical thinking using web-based argument mapping exercises with automated feedback. *Australasian Journal of Educational Technology*, 25(2), 268–291.
- Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need or cognition. *Journal of Personality Assessment*, 48, 306–307.
- Dawson, T. L. (2008). *Metacognition and learning in adulthood*. Northampton: Developmental Testing Service, LLC.
- Dwyer, C. P., Hogan, M. J., & Stewart, I. (2010). The evaluation of argument mapping as a learning tool: Comparing the effects of map reading versus text reading on comprehension and recall of arguments. *Thinking Skills and Creativity*, 5(1), 16–22.
- Dwyer, C. P., Hogan, M. J., & Stewart, I. (2011). *The promotion of critical thinking skills through argument mapping*. Nova Publishing, in press.
- Engelmann, T., & Hesse, F. W. (2010). How digital concept maps about the collaborators' knowledge and information influence computer-supported collaborative problem solving. *Computer-Supported Collaborative Learning*, 5, 299–319.
- Engelmann, T., Baumeister, A., Dingel, A., & Hesse, F. W. (2010). The added value of communication in a CSCCL-scenario compared to just having access to the partners' knowledge and information. In J. Sánchez, A. Cañas, & J. D. Novak (Eds.), *Concept maps making learning meaningful: Proceedings of the 4th international conference on concept mapping*, 1 (pp. 377–384). Viña del Mar: University of Chile.
- Ennis, R. H. (1989). Critical thinking and subject specificity: Clarification and needed research. *Educational Researcher*, 18, 4–10.
- Ennis, R. H. (1996). *Critical thinking*. Upper Saddle River: Prentice-Hall.
- Ennis, R. H. (1998). Is critical thinking culturally biased? *Teaching Philosophy*, 21(1), 15–33.
- Facione, P.A. (1990). The Delphi report. Committee on pre-college philosophy. *American Philosophical Association*.
- Farrand, P., Hussain, F., & Hennessy, E. (2002). The efficacy of the 'mind map' study technique. *Medical Education*, 36, 426–431.
- Flavell, J. (1979). Metacognition and cognitive monitoring: a new area of psychological inquiry. *American Psychologist*, 34, 906–911.
- Gadzella, B.M. (1996). *Teaching and learning critical thinking skills* (ERIC ED 405 313), U.S. Department of Education.
- García, T., Pintrich, P. R., & Paul, R. (1992). Critical thinking and its relationship to motivation, learning strategies and classroom experience. *Paper presented at the 100th Annual Meeting of the American Psychological Association, Washington, DC, August 14–18*.
- Halpern, D. F. (2003). *Thought & knowledge: An introduction to critical thinking* (4th ed.). New Jersey: Laurence Erlbaum Associates.
- Halpern, D. F. (2006). Is intelligence critical thinking? Why we need a new definition of intelligence. In P. C. Kyllonen, R. D. Roberts, & L. Stankov (Eds.), *Extending intelligence: Enhancement and new constructs* (pp. 293–310). New York: Taylor & Francis Group.
- Halpern, D. F. (2010). *The Halpern critical thinking assessment: Manual*. Vienna: Schuhfried.
- Harrell, M. (2004). The improvement of critical thinking skills. In *What philosophy is* (Tech. Rep. CMU-PHIL-158). Carnegie Mellon University, Department of Philosophy.
- Harrell, M. (2005). Using argument diagramming software in the classroom. *Teaching Philosophy*, 28, 2. [www.hss.cmu.edu/philosophy/harrell/ArgumentDiagramsInClassroom.pdf](http://www.hss.cmu.edu/philosophy/harrell/ArgumentDiagramsInClassroom.pdf).

- Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: a meta-analysis. *Review of Educational Research*, 66(2), 99–136.
- Higher Education Quality Council, Quality Enhancement Group. (1996). *What are graduates? Clarifying the attributes of "graduateness"*. London: HEQC.
- Hitchcock, D. (2003). *The effectiveness of computer-assisted instruction in critical thinking*. Philosophy department. McMaster University.
- Holmes, J., & Clizbe, E. (1997). Facing the 21st century. *Business Education Forum*, 52(1), 33–35.
- Huffaker, D. A., & Calvert, S. L. (2003). The new science of learning: active learning, metacognition and transfer of knowledge in e-learning applications. *Journal of Educational Computing Research*, 29(3), 325–334.
- Hwang, G. J., Shi, Y. R., & Chu, H. C. (2011). A concept map approach to developing collaborative mindtools for context-aware ubiquitous learning. *British Journal of Educational Technology*, 42(5), 778–789.
- Jensen, L. L. (1998). *The role of need for cognition in the development of reflective judgment*. Unpublished PhD thesis, University of Denver, Colorado, USA.
- Jiang, Y., Olson, I. R., & Chun, M. M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning Memory and Cognition*, 26, 683–702. 1997.
- Johnson, D.W., Johnson, R.T., & Stanne, M.S. (2000). Cooperative learning methods: A meta-analysis. Retrieved 21/06/2011, from <http://www.cooperation.org/pages/cl-methods.html>.
- King, P. M., & Kitchener, K. S. (2002). The reflective judgment model: Twenty years of epistemic cognition. In B. K. Hofer & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 37–61). New Jersey: Lawrence Erlbaum Associates.
- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85, 363–394.
- Ku, K. Y. L. (2009). Assessing students' critical thinking performance: urging for measurements using multi-response format. *Thinking Skills and Creativity*, 4(1), 70–76.
- Ku, K. Y. L., & Ho, I. T. (2010a). Dispositional factors predicting Chinese students' critical thinking performance. *Personality and Individual Differences*, 48, 54–58.
- Ku, K. Y. L., & Ho, I. T. (2010b). Metacognitive strategies that enhance critical thinking. *Metacognition Learning*, 5, 251–267.
- Kuhn, D. (1991). *The skills of argument*. Cambridge: Cambridge University Press.
- Kuhn, D., Goh, W., Iordanou, K., & Shaenfield, D. (2008). Arguing on the computer: a microgenetic study of developing argument skills in a computer-supported environment. *Child Development*, 79(5), 1310–1328.
- Ma, A. W. W. (2009). Computer supported collaborative learning and higher order thinking skills: A case study of textile studies. *The Interdisciplinary Journal of e-Learning and Learning Objects*, 5, 145–167.
- Marzano, R.J. (1998). *A theory-based meta-analysis of research on instruction*. Aurora, CO: Mid-Continent Regional Educational Laboratory. Retrieved 26/10/2007, from [http://www.mcrel.org/pdf/instruction/5982rr\\_instructionmeta\\_analysis.pdf](http://www.mcrel.org/pdf/instruction/5982rr_instructionmeta_analysis.pdf).
- Maybery, M. T., Bain, J. D., & Halford, G. S. (1986). Information-processing demands of transitive inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 12(4), 600–613.
- Mayer, R. E. (1997). Multimedia learning: are we asking the right questions? *Educational Psychologist*, 32(1), 1–19.
- Mayer, R. E. (2003). The promise of multimedia learning: using the same instructional design methods across different media. *Learning and Instruction*, 13, 125–139.
- Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63, 814–897.
- Monk, P. (2001). Mapping the future of argument. *Australian Financial Review*. 16, March, pp.8–9.
- National Academy of Sciences, National Academy of Engineering, Institute of Medicine. (2005). *Rising above the gathering storm: Energising and employing America for a brighter economic future*. Washington: Committee on Prospering in the Global Economy for the 21st Century.
- Norris, S. P. (1994). *The meaning of critical thinking test performance: The effects of abilities and dispositions on scores. Critical thinking: Current research, theory, and practice*. Dordrecht: Kluwer.
- Paivio, A. (1971). *Imagery and verbal processes*. Hillsdale: Erlbaum.
- Paivio, A. (1986). *Mental representations: A dual-coding approach*. New York: Oxford University Press.
- Paul, R. (1987). Dialogical thinking: Critical thought essential to the acquisition of rational knowledge and passions. In J. Baron & R. J. Sternberg (Eds.), *Teaching thinking skills: Theory and practice* (pp. 127–148). New York: W.H. Freeman.
- Paul, R. (1993). *Critical thinking: What every person needs to survive in a rapidly changing world*. Rohnert Park: Foundation for Critical Thinking.

- Perkins, D. N., Jay, E., & Tishman, S. (1993). Beyond abilities: a dispositional theory of thinking. *Merrill Palmer Quarterly*, 39, 1–1.
- 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)*. Michigan: National Center for Research to Improve Post-secondary Teaching and Learning.
- Reed, J. H., & Kromrey, J. D. (2001). Teaching critical thinking in a community college history course: empirical evidence from infusing Paul's model. *College Student Journal*, 35(2), 201–215.
- Robbins, S., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, 130(2), 261–288.
- Roth, W. M., & Roychoudhury, A. (1994). Science discourse through collaborative concept mapping: New perspectives for the teacher. *International Journal of Science Education*, 16, 437–455.
- Sherrard, M., & Czaja, R. (1999). Extending two cognitive processing scales: Need for cognition and need for evaluation for use in a health intervention. *European Advances in Consumer Research*, 4, 135–142.
- Solon, T. (2007). Generic critical thinking infusion and course content learning in introductory psychology. *Journal of Instructional Psychology*, 34(2), 95–109.
- Sweller, J. (1988). Cognitive load during problem solving: effects on learning. *Cognitive Science*, 12, 257–285.
- Sweller, J. (1999). *Instructional design in technical areas*. Australian Education Review No. 43. Victoria: Acer Press.
- Sweller, J. (2010). Cognitive load theory: Recent theoretical advances. In J. L. Plass, R. Moreno, & R. Brünken (Eds.), *Cognitive load theory* (pp. 29–47). New York: Cambridge University Press.
- Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied*, 3(4), 257–287.
- Toplak, M. E., & Stanovich, K. E. (2002). The domain specificity and generality of disjunctive reasoning: searching for a generalizable critical thinking skill. *Journal of Educational Psychology*, 94(1), 197–209.
- Twardy, C. R. (2004). Argument maps improve critical thinking. *Teaching Philosophy*, 27(2), 95–116.
- Valenzuela, J., Nieto, A. M., & Saiz, C. (2011). Critical thinking motivational scale: a contribution to the study of relationship between critical thinking and motivation. *Journal of Research in Educational Psychology*, 9(2), 823–848.
- van Eemeren, F. H., Grootendorst, R., Henkemans, F. S., Blair, J. A., Johnson, R. H., Krabbe, E. C. W., Planitin, C., Walton, D. N., Willard, C. A., Woods, J., & Zarefsky, D. (1996). *Fundamentals of argumentation theory: A handbook of historical backgrounds and contemporary developments*. New Jersey: Lawrence Erlbaum Associates.
- van Gelder, T. J. (2001). How to improve critical thinking using educational technology. In G. Kennedy, M. Keppell, C. McNaught, & T. Petrovic (Eds.), *Meeting at the crossroads: Proceedings of the 18th annual conference of the Australian society for computers in learning in tertiary education* (pp. 539–548). Melbourne: Biomedical Multimedia Unit, University of Melbourne.
- van Gelder, T. J. (2003). Enhancing deliberation through computer supported argument mapping. In P. Kirschner, S. Buckingham Shum, & C. Carr (Eds.), *Visualizing argumentation: Software tools for collaborative and educational sense-making* (pp. 97–115). London: Springer.
- van Gelder, T. J. (2007). The rationale for RationaleTM. *Law, Probability & Risk*, 6, 23–42.
- van Gelder, T. J., & Rizzo, A. (2001). Reason!Able across curriculum, in *Is IT an Odyssey in Learning? Proceedings of the 2001 Conference of ICT in Education*, Victoria, Australia.
- van Gelder, T. J., Bissett, M., & Cumming, G. (2004). Enhancing expertise in informal reasoning. *Canadian Journal of Experimental Psychology*, 58, 142–152.
- Wankat, P. (2002). *The effective efficient professor: Teaching, scholarship and service*. Boston: Allyn and Bacon.
- Wegerif, R. (2002). *Literature review in thinking skills, technology and learning: Report 2*. Bristol: NESTA Futurelab.
- Wegerif, R., & Dawes, L. (2004). *Thinking and learning with ICT: Raising achievement in the primary classroom*. London: Routledge.
- Willingham, D. T. (2007). Critical thinking: why is it so hard to teach? *American Educator*, 3, 8–19.
- Woodman, G. F., Vecera, S. P., & Luck, S. J. (2003). Perceptual organization influences visual working memory. *Psychonomic Bulletin & Review*, 10(1), 80–87.