

Obstacles and opportunities: addressing the growing pains of summative student evaluation of teaching

P.W.G. Surgenor

To cite this article: P.W.G. Surgenor (2013) Obstacles and opportunities: addressing the growing pains of summative student evaluation of teaching, *Assessment & Evaluation in Higher Education*, 38:3, 363-376, DOI: [10.1080/02602938.2011.635247](https://doi.org/10.1080/02602938.2011.635247)

To link to this article: <http://dx.doi.org/10.1080/02602938.2011.635247>



Published online: 16 Nov 2011.



Submit your article to this journal [↗](#)



Article views: 382



View related articles [↗](#)



Citing articles: 2 View citing articles [↗](#)

Obstacles and opportunities: addressing the growing pains of summative student evaluation of teaching

P.W.G. Surgenor*

UCD Teaching and Learning, University College Dublin, Dublin, Ireland

Summative student evaluation of teaching (SET) is a contentious process, but given the increasing emphasis on quality and accountability, as well as national and international calls for centralised student feedback systems, is likely to become an inevitable aspect of teaching. This research aimed to clarify academics' attitudes to SET in a large Irish university to identify obstacles to be addressed and opportunities to be developed. Manual thematic analysis was conducted on the transcriptions from four focus groups and written responses to focus group questions. Four themes emerged, relating to the purpose, validity, usefulness and consequences of SET. Opportunities to be developed included student accountability for students and easily obtained feedback on modules and teaching. Obstacles to be addressed included: challenges to academics' autonomy; speculation on potential negative consequences; difficulties measuring good teaching; and implications of SET data in both a personal and professional capacity. The research concludes by suggesting a reconsideration of the role of, and attitudes towards, SET in light of a shifting educational paradigm, and stating that SET has the potential to become an indicator of dedication to teaching and improvement, to ensure student feedback plays a role in promoting quality learning and not just quality stats.

Keywords: student evaluation of teaching (SET); summative; accountability; quality; opportunities

Introduction

Student evaluation of teaching (SET) when used summatively as a terminal indicator of mastery (Kealey 2010) is a contentious issue among academics (Beran and Rokosh 2009), many of whom consider it a superfluous and detrimental process with no real role in higher education (HE) (Sproule 2002). However, given the changing political and educational landscapes and associated demands for evidence of quality in teaching, summative SET is becoming an increasingly required fixture in European universities and in Ireland in particular. In addition to the reforms advocated in the Bologna process, the National Strategy Group on Higher Education (2011) recently delivered a report (commonly referred to as the Hunt Report) for the strategic national development for HE in Ireland, and recommended that Irish universities introduce regular summative centralised student feedback systems.

*Email: paul.surgenor@ucd.ie

This research contributes to existing literature by seeking to clarify lecturers' attitudes towards summative SET in Ireland in a changing educational landscape by identifying exactly which issues they perceive as obstacles or opportunities, and positing suggestions to address and maximise these issues, respectively.

The differing roles of SET

When used formatively, SET is an important and invaluable tool for academic development and has been universally accepted as such. It creates a mutually beneficial, personal channel of communication between academics and students (Zabaleta 2007), and demonstrates a lecturer's commitment to improving the students' learning experience. It is limited in value as an objective, independent measure of good teaching due to the lack of standardised formative measures, and provides no formal degree of accountability.

When used summatively SET data are collected centrally via institutionally defined or standardised surveys. They provide a quick, low-cost mechanism (Gray and Bergman 2003) designed to deliver a quantitative snapshot of a module or programme at a given time. The information generated tends to be less specific, providing a cursory overview of student responses to aspects of teaching and course development and delivery, and is therefore less beneficial for lecturers' pedagogic development than data from formative SET.

A review of current literature suggests that one of the major concerns of summative SET relates to the validity of the process. While the quantification of quality teaching is highly complex (Murphy, MacLaren, and Flynn 2009), many universities use student evaluation forms that have only the lowest level of accuracy and content validity (Hinton 1993; Yunker and Yunker 2003). Even well-established standardised forms have resulted in skewed and invalid data, casting doubt on the ability to measure quality teaching with precision (Wolfer and Johnson 2003). Suggested reasons for this difficulty include the abstract and subjective nature of good teaching (Cohen 1981), the reliance on overly simplistic and inflexible tools that fail to adequately reflect the complexity of the concept, and the (in)ability of students to objectively evaluate teaching (Beran and Rokosh 2009; d'Apollonia and Abrami 1997).

Furthermore, the list of confounding factors that may inadvertently influence students' comments is alarming long and varied and includes: age (Sproule 2002); gender of student (Heck, Todd, and Finn 2002); class size (Light and Cox 2001); student level of motivation and prior experience or ability (Langbein 2008); distribution of chocolate (Youmans and Jee 2007); perceived sexiness of the professor (Felton, Mitchell, and Stinson 2004); and use of humour (Adamson, O'Kane, and Shevlin 2005).

The need for quality and accountability at an institutional level has an inevitable impact on individual lecturers, who come under increasing pressure to demonstrate success at a modular level. This increasing demand for higher SET scores has been linked with grade inflation, the tendency to award higher grades in the hope of higher student ratings (Aleamoni 1999; Wachtel 1998), and is greater in institutions most dependent on tuition revenues (Langbein 2008). This phenomenon has been noted in Irish universities, where a recent report has suggested a 900% increase in the number of first-class honours degrees awarded between 2004 and 2008 (Walshe 2010).

Context

Unlike most North American universities, where mandatory summative SET has been largely accepted and implemented as a matter of course, adoption of centralised SET in European institutions has been slower, and most Irish universities at the time of this research had no mandatory centralised SET systems. While there have been reports of academics' reluctance to accept institutional-level feedback systems (Whitworth, Price, and Randall 2002), Irish lecturers will have limited opportunities to resist this inevitable change in SET practice for two main reasons.

National factors

More students are entering and remaining in HE in Ireland. There was a 50% increase in figures entering HE in Ireland in the last decade (Higher Education Authority 2009), during which time drop-out rates decreased (Higher Education Authority 2010; Higher Education Statistics Agency 2010). Education has been identified as an economic driver at a national and institutional level, as evidenced by the increased focus in attracting international students and promulgation of the country as a viable international figure in teaching quality and research (High-Level Group on International Education 2010). The success of such a gambit depends on the ability of Irish universities to demonstrate excellence not only in research, but also in teaching, which has traditionally been more difficult to objectively quantify (Langbein 2008). Consequently, and in line with the call for institution-level feedback systems in the Bologna Declaration, one recommendation from the recent report of the National Strategy Group on Higher Education (2011, 50) stated that 'Higher Education institutions should put in place systems to capture feedback from students, and use this feedback to inform institutional and programme management'. Regardless of any potential reticence, the introduction of summative SET to Irish universities is an inevitable component of demonstrating commitment to teaching quality.

Accountability

Sawbridge (1996) stated that the hierarchy found in most private and public sector enterprises is not evident in universities, and that the subsequent high degree of autonomy enjoyed by academics and the associated collegiality this engenders may be a barrier to more centralised control systems found in other organisations. Land (2001) recounts Sawbridge's (1996) discussion of the intellectual free trade of ideas creating a collegiality that requires academics to respect each other's intellectual independence. This results in a bottom-up approach to authority that is at odds with the newer type of managerialism increasingly evident in HE (Arthur 2009). The prevalence of employer initiatives, such as the introduction of centralised SET, can be regarded as contrary to the existing university culture (Sawbridge 1996) because of the increasing need for some degree of accountability.

In response to such claims, it could be argued that the current changes in the educational landscape exceeded notions of spurious 'employer initiatives' and questions the nature of accountability in academia. While academic freedom is central to HE and is a major strength of the system (High-Level Group on International Education 2010), Lu and Zhao (2010) stated that there was a responsibility to provide assurances of the quality of the teaching and learning experiences encountered by students. In the field of research while most academics are free to pursue

investigative lines of their own choosing they are, nevertheless, accountable to bodies such as ethics committees and peer-review panels to ensure quality. Similarly in the realm of teaching, while a lecturer should be free to engage in teaching practices and methods they feel most appropriate to their subject, there should be provision to ensure quality, and as the other party involved in this learning exchange, students should have a degree of input into this process. Such provision goes beyond the notion of fulfilling what Johnson (2000) calls a 'nominal bureaucratic function', providing instead some form of quantifiable indicator of teaching quality as found at primary and secondary educational levels.

Aims

Beran and Rokosh (2009) reported that while some instructors endorsed the idea of summative SET a significant proportion also gave strong negative responses. Given the contradictory nature of this area and the inevitability of institutional SET at Irish universities, this research aims to clarify this position with reference to a large Irish university and to identify any potential obstacles that should be addressed and opportunities that should be developed. This fulfils Murphy, MacLaren, and Flynn (2009) supposition that a culture of open discussion is an essential precursor to facilitate engagement and understanding and echoes a sentiment from Land (2001, 10) that states it is 'more effective to try and remove barriers if you can identify and locate them than just banging on trying to change things and hitting a brick wall all the time'.

Information generated will contribute to understanding academics' attitudes in a time of paradigm change, and provide insights that may make implementation a more tenable experience in the current university and in any European higher education institution facing the same imposition.

Method

Research was conducted within one of the largest universities in Ireland (approximately 25,000 students) that did not have a formal policy on summative SET. The university is divided into five colleges (C1–C5), and each college is comprised of a number of schools (two schools in the smallest college and 10 schools in the largest).

In order to ensure a representative sample, the school heads of teaching and learning in each college were contacted and informed of the study. They were provided with information and asked to discuss this with lecturers at two levels: within their school at their next school-level meeting to ensure that the views of all lecturers were made known to their school head; and secondly, with other school heads at a college level, so that an overview of the opinions of schools not attending the focus groups could be expressed to those who could attend.

A total of four focus groups were conducted with school heads of teaching and learning over a two-week period, one each for colleges C1–C4. Numbers participating in these groups ranged from four in the smallest to eight in the largest. A total of 25 school heads of teaching and learning participated in focus groups, representing not only their own views, but the views of the lecturers within their schools. In C5 the school heads of teaching and learning declined the opportunity to participate in a focus group, preferring instead to provide written responses to the focus group

questions. These were incorporated into the transcribed focus group responses for analysis.

Focus group questions were designed to spark discussion and to explore experiences and expectations of summative SET within their own college and at a university level. A sample of the questions included:

- What do you think is the main purpose of summative SET?
- What do you think are the barriers to conducting this type of SET?
- What implications would this type of SET have for your school?
- What do you think are the potential advantages and disadvantages of implementing a university-wide, formal SET policy?

Discussions were recorded and transcribed for manual thematic analysis. This method focuses on identifiable themes and patterns of living and/or behaviour (Aronson 1994) and involves identifying and coding data into meaningful units, which are reviewed, revised and grouped by theme.

The application of validity in the quantitative sense to qualitative research has traditionally proven problematic (Golafshani 2003), with doubt cast on even widely accepted qualitative practices such as triangulation and respondent validation (Silverman 2000). In this study a qualifying check was provided by engaging in a process of critical rationalism (Popper 1959). This challenges the potential for anecdotalism by requiring the researcher to refute their initial assumptions about their data in order to achieve objectivity and therefore temper the inclination to accept ostensible conclusions.

The anonymity of participants' responses was guaranteed and they were informed that they could pass on any question or remove themselves from the research at any stage.

Results

This research aimed to ascertain attitudes of academic staff towards compulsory, summative SET. Analysis resulted in the emergence of four SET-related themes: purpose; validity; usefulness of data and consequences. A summary of the key points for each theme is presented in Figures 1–4. In each figure the size of the area containing perceived obstacles and opportunities represents the proportion of positive and negative comments, illustrating the general attitude towards each issue at a glance.

Theme 1: purpose

The first emerging theme relates to the purpose of summative SET, and specifically to the reason for its introduction. As displayed in Figure 1, and contrary to expectation, there were more comments relating to opportunities than obstacles. Initial discussion expressed scepticism about its purpose, particularly its potential as a tool for surveillance:

... it would be for management to get information on lecturers and to penalise bad lecturers and reward good ones. (C1)

There was a suggestion that the process would be used to 'name and shame' lecturers identified to be unsuccessful or underperforming and that this may influence future

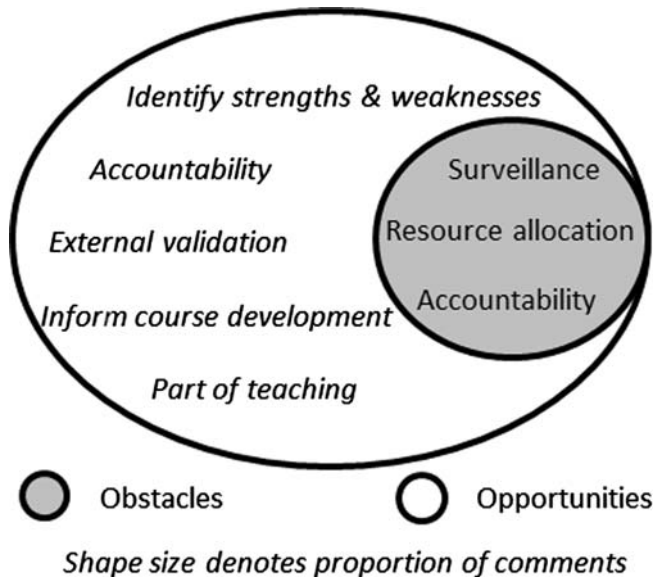


Figure 1. Summary of key obstacles and opportunities identified for Theme 1 (Purpose).

resource allocation. The introduction of a formal aspect of accountability was seen as a challenge to their current autonomy, and uncertainty over the purpose the process, access to the data, and its subsequent use were considered further obstacles:

... once the information is centralised how [is] that information going to be used and how [is it] going to be transferred once it's generated, how it flows from one place to another and whether or not schools or individual units would have the information that they originally submitted themselves. (C2)

The issue of accountability was identified as both an opportunity and an obstacle. A small proportion of lecturers stated it was their responsibility to ensure that checks and procedures for quality learning were in place and that this was part of their teaching remit. To this end they drew comparisons between the stringent requirements for research with the non-existent practices for teaching, while dissenters queried the ability to objectively define criteria for good teaching.

SET feedback was predicted to be a useful tool in the identification of strengths and weaknesses of their teaching practice and it was suggested that such independent and externally gathered data would be of use in validating promotion or tenure applications.

Theme 2: validity

The proportion of obstacles to opportunities illustrated in Figure 2 suggests the extent of apprehension related to issues of validity. While there was some positive sentiment about the possibility of a centralised system providing a uniform evaluation tool, and the degree of objectivity this would provide over lecturers' own non-standardised questionnaires, participants predominantly regarded validity-related concerns to be a considerable obstacle to SET.

The perceived administrative nature of summative SET prompted speculation on its relevance to pedagogy and for their teaching practice, and it was suggested that lecturers should not be forced to participate in an irrelevant process that detracted from their teaching duties:

It's a total waste of time (C1) . . . [lecturers] shouldn't be subjected to that kind of process (C2) . . . there are a variety of attitudes among the staff about teaching and what that means. (C3)

There was debate as to whether any evaluation tool was capable of generating valid data that would, as one participant claimed, distinguish between student perceptions of good teaching and difficult material:

. . . there's no one can teach it in such a way that [doesn't require] putting in the work. If it's not easy, it's not easy – therefore they think it must be poorly taught. (C4)

This obstacle was further complicated by concerns of sampling bias inherent in the process since end of term SET invariably drew on either keen, committed students or dissatisfied students, resulting in comments that were glowing praise or vitriolic scorn, with little middle ground. The SET process also attributed equal weighting to comments from regular attending non-regular attending students:

You have all sorts of people filling them in who don't attend lectures. Sometimes the responses you get to student questionnaires are so off the mark that you know the students haven't been at the lectures (C1) . . . there would be some students evaluating modules they hadn't attended, or attended only in part. And that's a very big concern. (C2)

While this would be beneficial for formative SET to understand the reasons for non-attendance, the impact of such students on the quantitative mean scores that may impact on a lecturer's professional career development was called into question.

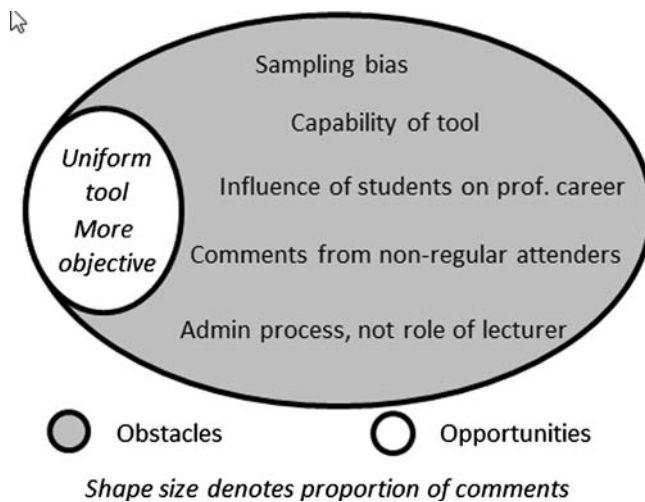


Figure 2. Summary of key obstacles and opportunities identified for Theme 2 (Validity).

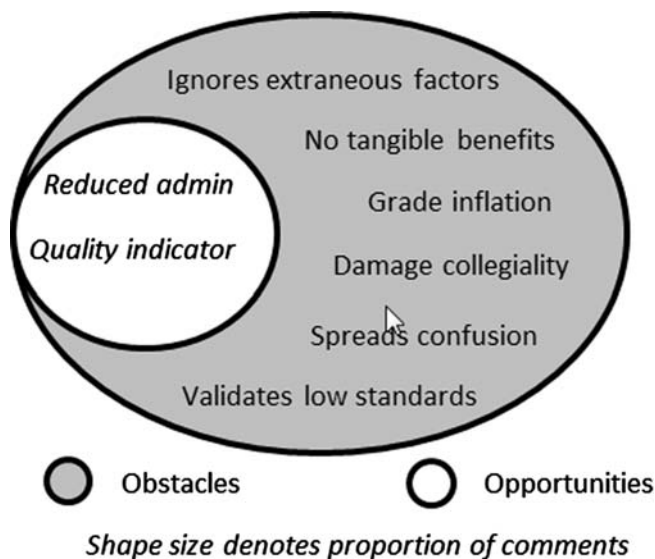


Figure 3. Summary of key obstacles and opportunities identified for Theme 3 (Usefulness).

Theme 3: usefulness of data

There were more perceived obstacles than opportunities in relation to the usefulness of data generated by summative SET. Standardised quantitative questionnaires were not believed to be capable of generating tangible benefits in terms of constructive feedback for personal improvement or development. Conversely, it was feared that SET data may be detrimental by spreading confusion rather than clarification:

I don't think it generates meaningful data for our practice as lecturers, as teachers. And I don't think it helps the world outside in terms of understanding what it is we're doing. It suggests that things are comparable when they're not. It gives people quite superficial feedback that doesn't help them to make decisions. (C2)

The use of average figures from a generic evaluation tool was seen to offer little in terms of direction, but may have the unintentional effect of validating low standards and damaging collegiality:

The people who were up very high thought it was good, but all other people felt a bit aggrieved, and there was this whole thing, well it's alright for him, he's doing an easy course, or maybe he's doing his exams, you know – a whole lot of analysis that was absolutely useless. (C1)

This point also touches upon the issue of recognising extraneous variables that may influence SET scores, such as class size, time or difficulty. There was discussion of the possibility of tactics similar to grade inflation to ensure higher SET scores:

If [resource allocation] depends on it you can bet that we're going to try and game those questionnaires so we get the right results, and so are all of you if you're smart! (C5)

Theme 4: consequences

The consequences of summative evaluation were an unknown quantity and, as can be seen in Figure 4, focused more on possible negative outcomes. The first concern related to the perceived encroachment on their autonomy and element of professional criticism that the introduction of mandatory SET would inevitably bring:

... that's why a lot of staff don't buy in because they may know that they're likely to be criticised and they're wondering what use will be made of that information and how public it may become. (C4)

An associated issue related to speculation over data use. There was a strong degree of uncertainty over how the information would be used and lecturers were apprehensive about who would have access to this confidential, potentially sensitive data and the possibility for this to be exploited in some manner:

People feel very sensitive about this information, why we're gathering it and what it's being used for (C4) ... People would worry about that, in terms of how the data would actually be used (C2) ... If you have some particularly nasty individual they can use it in a pretty dodgy way. (C5)

Initial fears were exacerbated during discussions and in several cases developed into concerns about implications for low student ratings. In one instance a group discussed the possibility that poor performance would result in punishing the lecturer responsible and since no-one could either validate or extirpate the rumour, it persisted.

Not all of the consequences of SET were considered barriers, however, and several opportunities also emerged. Summative SET was believed to have the potential

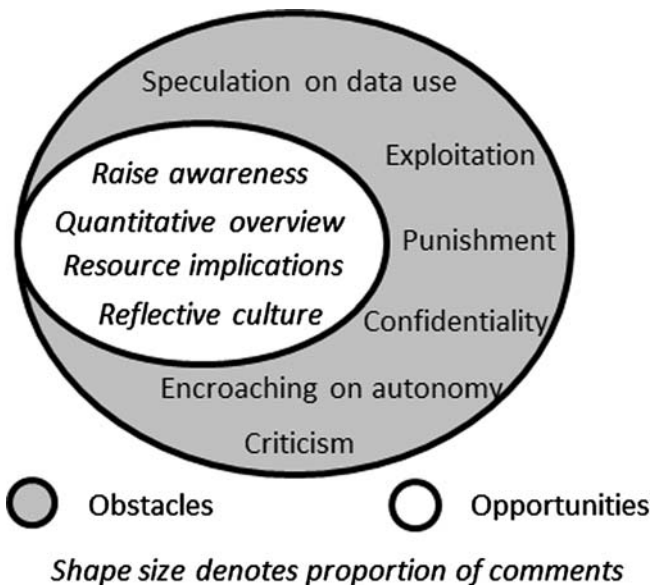


Figure 4. Summary of key obstacles and opportunities identified for Theme 4 (Consequences).

to raise general awareness of evaluation and ‘promote teaching within the university’ (C5). Although not as beneficial as formative SET, it was seen as an opportunity for academics to receive a quantitative overview of their teaching, including those who may not normally engage in any form of SET. Subsequently there was a sense this may lead to greater personal reflection on pedagogy and prove an important step in establishing an institutional culture of reflection and evaluation:

I’d see the positive aspect of it being university wide as being part of a culture, a reflection on teaching, a constructive reflection on teaching and what we’re doing.
(C2)

A standardised central SET system ‘would reduce a certain amount of the administrative load for module coordinators’ (C2) and, as well as saving on resources, would also remove the need for academics to navigate the often conflicting minefield of SET literature and instruments.

Summary and discussion

The aim of this study was to identify potential obstacles to be addressed, opportunities to be developed and any insights that would make implementation of summative SET feedback systems more tenable within this and other universities in the same position. Four themes emerged from the analysis of focus group data, with obstacles and opportunities identified in relation to the purpose, validity, usefulness and consequences of summative SET.

Opportunities afforded by SET related to students, modules and to lecturers themselves. SET would provide an unheralded level of accountability for all students, not just those whose lecturers instigated their own student feedback systems. If embraced, this should have a beneficial impact on student empowerment and sense of involvement in, and responsibility for, their education at a time when both they and society are demanding more from HE than just discipline-specific knowledge. Academics identified a number of potential benefits from engaging in centralised SET. Firstly a centralised summative evaluation system was perceived as an attractive prospect in relation to the administrative load traditionally associated with large scale evaluations of teaching. It would enable the prospect to gather evidence of teaching quality and a centralised channel of communication with students through which good teaching can be acknowledged. A centrally administered SET system was seen as an opportunity to provide a more objective, independent verification of teaching quality that could be used to develop some aspect of their module or as evidence in portfolios or promotional material.

Obstacles to summative SET can be grouped into four categories, relating to tradition, speculation, validation or implications (see Figure 5).

‘Tradition’ obstacles refer to the challenges that summative SET presents to lecturers’ traditional autonomy and level of accountability, and is closely associated with the changes effected by the increasing managerialist influence in HE (Arthur 2009). Such concerns represent the traditional paradigm of academia, unencumbered by the demands of transparency and accountability from an increasingly consumerist student body. The social and political changes represented in the recommendations from Bologna and the Irish National Strategy Group on Higher Education (2011) address this shifting paradigm by requiring an evaluation of established roles

and a greater accountability for the teaching and learning experiences facilitated by the lecturer. While academics may attempt to resist this evolution of our educational culture, given the increased demands for transparency and evidence of quality, such attempts are likely to be in vain.

‘Speculation’ obstacles are those likely to accompany the introduction of any new system before specific details are available, and may be addressed by the dissemination of clear, well-defined policies. One method to counter the spread of fears and suspicions is to ensure academic staff are aware of the process, and if possible, involved in its development. This engagement would afford opportunities to explore the potential for SET to improve teaching and collegiality through pedagogic discussion and the exchange of ideas rather than threaten it as initially feared.

‘Validation’ obstacles encompass the difficulties associated with measurement and validity of an essentially subjective concept like ‘good teaching’. The myriad issues with validity are well charted in SET literature, as are the numerous suggestions for navigating this potential minefield. By identifying, propagating and acceding to best practice in this area (e.g. multi-method evaluation, the coadjutant use formative and summative SET, collaboration of SET data from additional sources, support with analysing results and implementing changes where appropriate) many of these issues may be avoided. In terms of the SET tool, where possible, this should be based on established, validated standardised instrument that addresses the complex and multidimensional nature of teaching and learning. Edström (2008) suggests that surveys can have either a teacher or a learner focus, concerned with the teacher or the teaching process in the case of the former, and the learning outcomes and the learning process in the case of the latter. Surveys that focus on teaching not only have the potential to be highly influenced by the lecturer’s personality, but may reward more traditional disengaged transmission teaching methods rather than more engaging and cognitively demanding ones. Therefore, focusing on the learning process may be one way to circumvent some of the validation issues related with measurement. One suggestion that arose from discussion to avoid low-attendees

	Tradition	Speculation	Validation	Implications
Obstacle	Challenge to traditional autonomy and level of accountability	Confusion over role and output of SET	Issues of validity relating to tool and process	Impact of potentially flawed data at professional and personal levels
Solution	Accept the changing roles & responsibilities of academics	Clarification through discussion and transparent policies	Identify methodological issues & employ standardised tools	Development of research-informed policies for data use

Figure 5. Summary of obstacles to, and suggested solutions for, summative SET by category.

from skewing data (particularly if used in personnel decisions) was to include a measure on attendance and filter out those with near-zero attendance prior to any statistical analysis. While every student would still have the opportunity to comment, the ratings of those who had attended few or no classes would not adversely impact scores on items addressing classroom learning or interaction.

'Implications' obstacles incorporate issues resulting from the ineffective use of SET data, such as potential exploitation, punishment, the consequences of using non-representative data in promotion decisions, validating low teaching standards or actions leading to grade inflation. The fact that, for many, these were the most pressing issues illustrates a major problem with SET, i.e. that much concern expressed was not due to the evaluation process itself which, as illustrated in Figure 1, was greeted with a considerable degree of positivity. Rather, concerns focused on the potential misuse of SET data or the possibility it may misrepresent or inaccurately reflect their teaching. Due to the necessary brevity of SET instruments and the conversely expansive scope of teaching, summative SET is relatively ineffective as a standalone indicator of teaching quality. For this reason it may be necessary to review and redefine the process more as an indicator of dedication to teaching and improvement. This would require expanding the current role of SET, which frequently ends with data collection, to incorporate evidence of accumulating and amalgamating additional sources of information on teaching. This could also include evidence documenting reflection on suggested changes, reasons for implementing (or not implementing) such changes, and how subsequent developments are communicated to students, fellow lecturers or heads of school.

Conclusions

Land (2001) suggests that like other professional-groups academics have no interest in maintaining the status quo, a supposition supported by the fact that attitudes to summative SET in this study were not uniformly negative. There were, however, potential barriers associated with the period of change in which academics currently find themselves. To ensure their role does not become superfluous, academics must engage in this process and contribute to development. This paper proposes practical suggestions which may be of use to academics and Higher Education Institutions when introducing or reviewing the role of SET. Given the rise of managerialism and the national and international edicts on accountability and quality it is time to reconsider our attitude towards summative SET, to ensure that it plays a role in promoting quality learning and not just quality stats.

Notes on contributor

P.W.G. Surgenor is a lecturer in Educational Development at University College Dublin, Ireland. His interests include psychology in education, student evaluation of teaching, and the first year experience.

References

- Adamson, G., D. O'Kane, and M. Shevlin. 2005. Students' ratings of teaching effectiveness: A laughing matter? *Psychological Reports* 96, no. 1: 225–6.
- Aleamoni, L.M. 1999. Student rating myths versus research facts from 1924 to 1998. *Journal of Personnel Evaluation in Education* 13, no. 1: 153–66.

- Aronson, J. 1994. A pragmatic view of thematic analysis. *The Qualitative Report* 2, no. 1. <http://www.nova.edu/ssss/QR/BackIssues/QR2-1/aronson.html>.
- Arthur, L. 2009. From performativity to professionalism: Lecturers' responses to student feedback. *Teaching in Higher Education* 14, no. 4: 441–54.
- Beran, T.N., and J.L. Rokosh. 2009. Instructors' perspectives on the utility of student ratings of instruction. *Instructional Science* 37, no. 2: 171–84.
- Cohen, P.A. 1981. Student ratings of instruction and student achievement: A meta-analysis of multisection validity studies. *Research in Higher Education* 51, no. 3: 281–309.
- d'Apollonia, S., and P.C. Abrami. 1997. Navigating student ratings of instruction. *American Psychologist* 52: 1198–208.
- Edström, K. 2008. Doing course evaluation as if learning matters most. *Higher Education Research & Development* 27, no. 2: 95–106.
- Felton, J., J. Mitchell, and M. Stinson. 2004. Web-based student evaluation of professors: The relations between perceived quality, easiness and sexiness. *Assessment & Evaluation in Higher Education* 29, no. 1: 91–108.
- Golafshani, N. 2003. Understanding reliability and validity in qualitative research. *The Qualitative Report* 8, no. 4: 597–607.
- Gray, M., and B.R. Bergman. 2003. Student teaching evaluations: Inaccurate, demeaning, misused. *Academe* October 89, no. 5: 44–6.
- Heck, J., J. Todd, and D. Finn. 2002. Is student performance enhanced by perceived teaching quality? *Journal of Financial Education* 28: 54–62.
- High-Level Group on International Education. 2010. *Investing in global relationships Ireland's international education strategy 2010–2015*. Dublin: Department of Education and Skills.
- Higher Education Authority. 2009. Record numbers going to higher education. <http://www.heai.ie/en/node/1317>.
- Higher Education Authority. 2010. *A study of progression in higher education*. Dublin: HEA.
- Higher Education Statistics Agency. 2010. *Performance indicators in higher education in the UK 2008/09*. Cheltenham: HESA.
- Hinton, H. 1993. Reliability and validity of student evaluations: Testing models versus survey research models. *Political Science and Politics* 26, no. 3: 562–70.
- Johnson, R. 2000. The authority of the student evaluation questionnaire. *Teaching in Higher Education* 5, no. 4: 419–34.
- Kealey, E. 2010. Assessment and evaluation in social work education: Formative and summative approaches. *Journal of Teaching in Social Work* 30, no. 1: 64–74.
- Land, R. 2001. Agency, context, and change in academic development. *The International Journal for Academic Development* 6, no. 1: 5–20.
- Langbein, L. 2008. Management by results: Student evaluation of faculty teaching and the mis-measurement of performance. *Economics of Education Review* 27, no. 4: 417–28.
- Light, G., and R. Cox. 2001. *Learning and teaching in higher education: The reflective professional*. London: Sage.
- Lu, W., and S. Zhao. 2010. Development on quality assurance of teaching and learning. *Management Science and Engineering* 4, no. 2: 62–8.
- Murphy, T., I. MacLaren, and S. Flynn. 2009. Toward a summative system for the assessment of teaching quality in higher education. *International Journal of Teaching and Learning in Higher Education* 20, no. 2: 226–36.
- National Strategy Group on Higher Education. 2011. *National strategy for higher education to 2030: Report of the strategy group*. Dublin: Department of Education and Skills.
- Popper, K. 1959. *The logic of scientific discovery*. New York, NY: Basic Books.
- Sawbridge, M. 1996. The politics and organisational complexity of staff development for academics: A discussion paper. UCoSDA Occasional Green Paper, No. 14, Sheffield, UK.
- Silverman, D. 2000. *Doing qualitative research*. London: Sage.
- Sproule, R. 2002. The underdetermination of instructor performance by data from the student evaluation of teaching. *Economics of Education Review* 21: 287–94.
- Wachtel, H.K. 1998. Student evaluation of college teaching effectiveness: A brief review. *Assessment & Evaluation in Higher Education* 23, no. 2: 191–211.

- Walshe, J. 2010. O’Keeffe orders probe as college scores 900pc inflation in grades. *Irish Independent*, March 2. <http://www.independent.ie/education/latest-news/okeeffe-orders-probe-as-college-scores-900pc-inflation-in-grades-2085229.html>.
- Whitworth, J.E., B.A. Price, and C.H. Randall. 2002. Factors that affect business student opinion of teaching and learning. *Journal of Education for Business* 77, no. 5: 282–9.
- Wolfer, T., and M. Johnson. 2003. Re-evaluating student evaluation of teaching: The teaching evaluation form. *Journal of Social Work Education* 39: 111–21.
- Youmans, R.J., and B.D. Jee. 2007. Fudging the numbers: Distributing chocolate influences student evaluations of an undergraduate course. *Teaching of Psychology* 34, no. 4: 245–7.
- Yunker, P., and J. Yunker. 2003. Are student evaluations of teaching valid? Evidence from an analytical business core course. *Journal of Education for Business* 78, no. 6: 313–7.
- Zabaleta, F. 2007. The use and misuse of student evaluations of teaching. *Teaching in Higher Education* 12, no. 1: 55–76.