

Development of the Canterbury Christ Church University Module Evaluation Questionnaire: an overview of the research

Dr Susan Kenyon, Faculty Director of Learning and Teaching

January 2020

Susan.Kenyon@Canterbury.ac.uk

This paper summarises the research that was undertaken to develop the CCCU Module Evaluations Questionnaire (MEQ) during academic year 2018-2019. The MEQ was accepted as the standard MEQ, to be used by all Programmes, at the Education and Student Experience Committee in November 2018.

The author gratefully acknowledges the contribution of all staff and students who took part in this research.

The views expressed in this paper are those of the author and do not necessarily reflect the views of the University.

Abstract

Student evaluations of teaching (SETs) are a ubiquitous tool in the Higher Education tool box. SETs facilitate the student voice, enabling students to offer feedback, views and opinions on the curriculum and pedagogy, which can provide an invaluable evidence-base for enhancement and change. However, SETs suffer from a reputational problem. Academics question their reliability and validity; senior leaders question their utility in delivering and measuring institutional improvement; students are sceptical of their impact on teacher behaviour and their future learning. This paper reports research that sought to design a SET questionnaire that could meet the full potential of the method. To do this, the study sought to understand the varying needs that key actors in HE have from SET. The resulting questionnaire enables an unfettered student voice, providing qualitative evidence for enhancement, whilst simultaneously collecting quantitative data on key metrics to enable monitoring and a degree of accountability.

Introduction

Higher education (HE) exists to support learning, for the benefit of individuals (employability, wellbeing), communities (cohesion, civic participation, reduction in antisocial behaviour) and wider global society (economic growth, multi-cultural understanding, tolerance) (Chan, 2016; HEFCE, 2017; Kenyon, 2015).

We can evaluate how successful HE has been in supporting learning through assessment, measuring change in knowledge acquisition, skills application and understanding (Biggs and Tang, 2011). However, it is essential that we also evaluate the effectiveness of what happens in the classroom in supporting student learning. This helps to ensure that the quality of teaching is the best that it can be, providing evidence of what is working, which can be continued and shared as best practice, as well as evidence to inform teaching enhancement, where needed.

To do this, Higher Education Institutions (HEIs) commonly ask students to evaluate the effectiveness of the teaching that they have experienced in supporting their learning, via Student Evaluation of Teaching (SET) questionnaires (Spooren and Christiaens, 2017).

However, SET suffers from a reputational problem. Academics perceive student evaluations of their teaching with 'scepticism, uncertainty or unease' (Darwin, 2017: 14). Senior leaders question the utility of SET in providing data for the measurement of institutional improvement, comparison and change in the same (Kahu, 2013). Students are sceptical of the impact of SETs on teacher behaviour, on their future learning and on wider institutional behaviours (Spooren and Christiaens, 2017). This scepticism manifests itself in the very low response rates to SETs, across the sector (Alhija, 2017) which, in turn, compounds academic and senior leader scepticism of their validity and reliability. The research literature supports these concerns.

However, there is great potential in SET. SET facilitates the student voice, enabling students to offer feedback, views and opinions on the curriculum and pedagogy, which can provide an invaluable evidence-base for enhancement and change (Healey, Flint and Harrington, 2014; Zepke, 2018). SET can promote student engagement, with positive impacts on outcomes and learning (Kahu, 2013). It follows that harnessing this potential could have substantial benefit, for all stakeholders in HE.

This paper reports research that sought to design a SET questionnaire that could meet the full potential of the method. The research was conceived in response to a policy deadlock at CCCU. The University wished to introduce a standard module evaluation questionnaire (MEQ), to contribute to

the enhancement of learning, teaching and student outcomes. Multiple draft MEQs had been submitted over five years, but it had not been possible to agree a way forward.

The author sought to understand the reasons for this impasse. It emerged that underpinning the deadlock was an informed by an ultimately irresolvable debate regarding the purpose of SET. As a result, it had not been possible to agree a way forward with the design of an MEQ.

To resolve this, the study sought to understand the varying needs that key actors in HE have from SETs and to investigate whether or not it is possible to design a SET that meets the potential of the medium, delivering a valid, reliable impression of the student voice in a way that satisfies the different needs of academics, senior leaders and students.

The paper progresses as follows. First, a literature review highlights the key difficulties with and benefits of SETs. The methodology follows. The findings for each of the four Stages of research are presented, before the final MEQ design is revealed. A discussion section reflects on the potential of the MEQ, before concluding by acknowledging the limitations of this research and highlighting directions for future research.

Literature review

Student evaluations of teaching: three key concerns

Alhija (2017a) summarises three key concerns with student evaluations of teaching, which underlie the scepticism, uncertainty and unease reported by Darwin (op cit). The first of these relates to the *content and structure of the evaluation tools*. The aim of SETs is to improve teaching through feedback. However, for Alhija (ibid) and others, the tools used for SETs – namely, quantitative evaluation questionnaires – do not collect the right data for this. As such, they cannot provide feedback to improve teacher performance.

In 2006, Nicol and Macfarlane-Dick suggested seven principles of good feedback, which must be present if feedback is to be successful in improving student performance. The authors' first principle is that good feedback must clarify what good practice is. If we apply this principle to feedback via SETs, because there is little agreement on what 'good' teaching, in the research literature, between and within Institutions, between and within disciplines, even amongst academic practitioners on the same Programme, feedback via SETs is unlikely to help the teacher to improve their teaching.

In the absence of an agreed definition of the 'goals, criteria, expected standards' (ibid: 2015) that the teacher must work towards, SETs can increase in complexity, to encompass the multiple definitions and conceptualisations of 'good teaching'. Utriainen et al (2018) discuss the 77-question questionnaire 'Experiences of Teaching and Learning Questionnaire', which was designed for this purpose. However, the authors find that even this fails to measure teaching quality to their satisfaction.

Alternatively, SETs can focus on *student satisfaction*, rather than the quality of teaching and student learning. However, satisfaction and teaching quality are not necessarily linked. For example, in their meta-analysis of SET ratings and student achievement, Uttl et al (2017) the authors find no correlation between SET ratings and student learning.

Because SETs are not measuring what they need to measure, they are not providing effective feedback to improve teaching performance. Returning to Nicol and Macfarlane's definition of good feedback: they will not facilitate reflection; delivery high quality information to teachers about

their teaching; encourage dialogue between teachers and their students; encourage positive motivation, belief and esteem; or provide the opportunity to close the gap between current and desired teacher performance (2006: 205). It follows that they are unlikely to enhance student learning.

Furthermore, when SETs measure satisfaction, rather than teaching quality, there is *substantial potential for bias*, which further reduces the utility of feedback for teaching improvement – Alhija's second key concern (2017a).

Multiple authors highlight bias in student responses, suggesting that SETs do not do not perform consistently across different types of students and staff. For example, Darwin highlights the perception amongst academics that SETs are inappropriate for our 'ever more heterogeneous, technologically-immersed and demanding student populations' (2017: 14), raising the question of the reliability of standardised SETs within an increasingly diverse student community. Alhija (2017b) highlights differential judgement by students with different characteristics. Spooren and Christiaens (2017) report that SET scores are significantly influenced by students' perceptions of the influence and value of SETs. And MacNell et al (2015) report a significant gender bias in SETs, with students rating male teachers significantly higher than female teachers; a finding supported by studies by Mitchell and Martin (2018) and by Boring (2017), who finds no relationship between gender and teaching effectiveness when independently measured, despite lower SET ratings for teaching effectiveness for female teachers.

Alhija's third key concern is the *misuse of SET results*, that is, for reasons other than pedagogical improvement (2017a). Kahu reminds us that many SETs are designed as tools for institutional improvement and comparison, not learning and teaching enhancement (2013: 760). Other authors suggest that SETs are used to justify administrative decisions, such as Programme closure, personnel decisions, including appointment, promotion, performance management (Linse, 2017; Nielsen and Kreiner, 2017). In the UK, SET is used by central government to inform the Teaching Excellence and Student Outcomes Framework (TEF), which ranks Universities according to 'teaching excellence'. Stated motivations for the TEF include the student-focused – to drive up the standards of teaching and encourage Universities to work with students to enhance the student experience. However, the TEF also aims to enhance the market in Higher Education, using former students' satisfaction with teaching to influence future students' choices; linking SETs with Programme fees; and using SETs assess whether or not students have had a high quality academic experience that is worthy of the time and money that they have spent in Higher Education (HE), inextricably linking SET to Quality Assurance and financial accountability, rather than pedagogical improvement (OfS, 2018).

The brief review, above, appears to condemn SETs as unreliable, invalid, ineffective tools in the University teacher's toolbox. Darwin goes further, suggesting that SETs may have a negative effect on teaching quality, 'encourag[ing] teachers to adopt pedagogical or assessment strategies that appease student opinion, resulting in the danger of diminishing academic standards as a compensatory outcome for improving student ratings' (2017: 21). Zepke (2018) suggests that the narrow focus of SETs simplifies and generalises the 'messy reality' and complexity of learning and teaching. Citing Howie and Bagnall, Zepke further suggests that the overriding purpose of SETs is simply to confirm existing ideas; that Institutional control of the same can reduce the student voice to a tool of the Institution, in the service of government, to justify HE policy. Recent changes to the TEF in the UK to increase the weight of the student voice in Institutional rankings could be perceived to reinforce this view (OfS, 2018b).

Student evaluations of teaching: the case for SET

However, there is great value in hearing the student voice. Healey, Flint and Harrington (2014) describe the benefits of the student voice for facilitating learning and teaching enhancement, with benefits for student attainment, community, engagement, inclusion, retention and staff satisfaction. For these authors, SET is fundamental in facilitating the student voice, enabling students to offer feedback, views and opinions on the curriculum and pedagogy, which provides an invaluable evidence-base for enhancement and change (ibid: ch. 3). SETs also develop students' skills in reflection and evaluation, which are invaluable graduate attributes (Zepke, 2015).

For others, the importance of the student voice is broader than learning and teaching enhancement. Incorporating the student feedback speaks to principles of democracy in education by levelling the hierarchy between students and lecturers. It challenges 'dominant notions of students as customers and education as a commodity' (Matthews et al., 2018: 9), transforming students from consumers to co-producers of their learning. When we facilitate the student voice, we promote 'active student citizenship' and 'students' cognitive investment in and emotional commitment to their learning', each with substantial individual and social benefits (Zepke, 2018: 437, 438).

Given the improvement potential offered by the student voice, it is essential to design a way of capturing the student voice that realises the benefits of student participation in the design and delivery of their learning and teaching. But is it possible for SETs to achieve this, whilst avoiding the key concerns expressed by academics, senior leaders and students, outlined above?

The study below aimed to achieve this. This paper now turns to discuss the study, presenting the methodology, before considering the findings from each Stage of the research.

Method

The method was designed to investigate three key actors/stakeholders' perceptions of SET in HE: academics; senior leaders of learning and teaching; and students. The aim was to understand each actors' needs from SET and then to develop an MEQ in partnership, which meets all needs, to overcome policy deadlock and realise the potential of SETs to facilitate the student voice and enable students to be partners in their learning.

This research employed an exploratory sequential design approach (Creswell and Plano Clark, 2007), combining qualitative and quantitative research over four Stages. Research design was emergent. The results from Stage 1 were used to determine the research questions, methods and population for Stage 2; results from Stage 2 determined Stage 3; and Stage 3 determined Stage 4. Data analysis was inductive and followed a modified grounded theory approach: themes were developed through the process of analysis and the theories proposed in the results section were emergent from these themes.

Figure 1 summarises the methodology, highlighting the depth and breadth of this research. Stage 3, which involved students, received full ethical clearance from the Faculty of Social and Applied Sciences Ethics Review Committee.

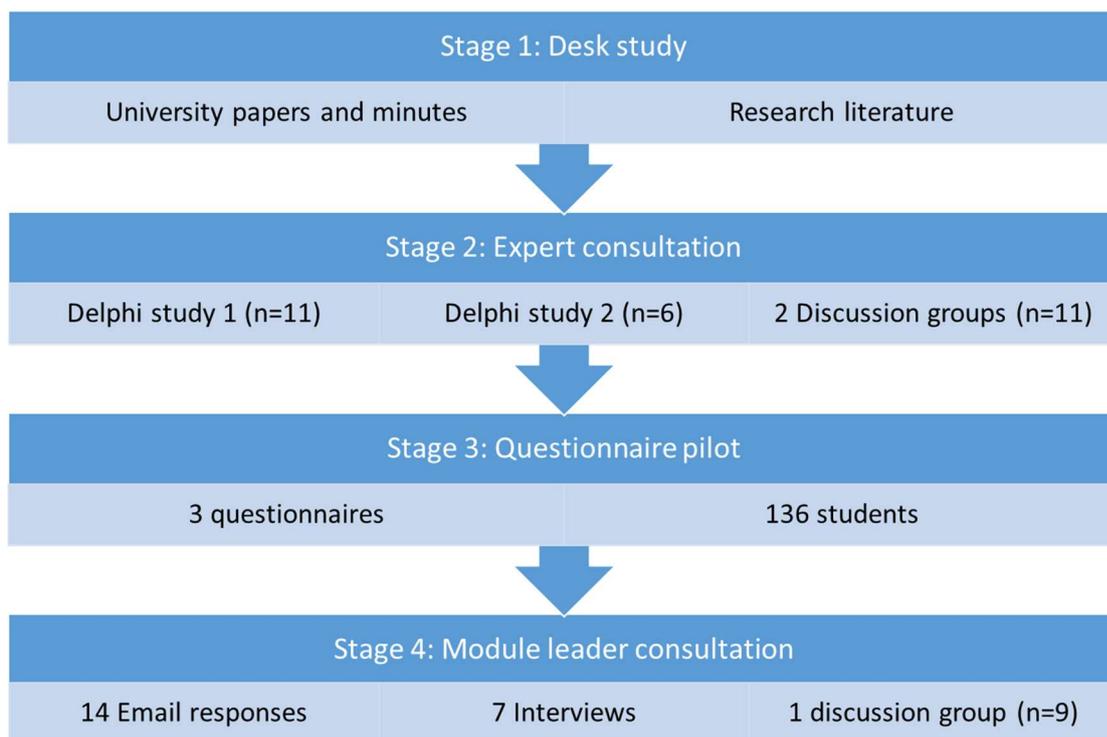


Figure 1. Methodology.

This paper now turns to consider the results from each Stage of the research.

Results

Stage 1: Desk study

The first stage of the research aimed to understand module evaluations policy to date. This involved a desk study of papers discussing module evaluations, submitted to University Committees from 2014 onwards, plus minutes from these meetings.

Analysis of proposed policy papers, alongside recorded accounts of these discussions, revealed a clear ‘tension between improvement and accountability motives’ for the new MEQ (Darwin, 2017: 13). It emerged strongly that different actors at the Institution had different visions of the purpose of SET. Whilst all actors focussed on improvement, the perception of what improvement looks like, alongside perceptions of the way to drive improvement, differed.

Broadly, the analysis suggested a deep division between MEQs as a driver for teaching enhancement and MEQs as a tool for monitoring satisfaction with teaching. Following from this, the perceptions of the type of data that should be collected in an MEQ differed. For the former group, MEQs should record how students feel about the teaching that they have experienced. For the latter group, MEQs should provide quantitative ratings of the teaching that students have received.

Over time, these positions appeared to become more polarised, as actors became more convinced of their initial position (a phenomenon known as ‘attitude polarisation’ (Lord et al, 1979)). Consensus appeared to reduce through discussion and policy progress had stalled.

Stage 1 revealed that the key reason underlying lack of progress in introducing a standard MEQ was this tension between the different demands from MEQs. This finding confirmed the findings from the literature review and informed the next Stage of the research. Stage 2 was designed to enable the expression of different views, to further understanding of different actors' needs, but to do this in a controlled way, to minimise further attitude polarisation. Stage 2 further aimed to pull colleagues towards consensus, if possible, exploring if MEQs should gather data for enhancement/improvement, if they should gather data for monitoring/accountability, or if they could do both.

Stage 2: Expert consultation

Two online Delphi studies were conducted, followed by two discussion groups. Delphi studies were chosen to enable actors to contribute their views equally, but in isolation. The author, as facilitator, examined responses for consensus and only areas of consensus were taken forward to the next study. This minimised the opportunity for further attitude polarisation and allowed the consultation to be more inclusive and less combative than on conducted in real-time, or face-to-face.

Stage 2 included representatives from all Faculties, Central Services and the Students Union. It included academics and professional services staff. 15 were invited to participate in the Delphi studies. 11 participated in Delphi 1. 6 participated in Delphi 2. All who participated in the Delphi studies were invited to take part in a discussion group. 12 were invited and 11 participated.

Stage 2 confirmed the tension between improvement/enhancement and monitoring/accountability. It emerged strongly that participants were divided largely along the lines of their job role. The majority of senior leaders of learning and teaching expressed the view that improvement is driven by metrics. Metrics enable benchmarking, which enables us to gauge relative success and change in the same. Therefore, for most of these participants, SETs must provide data that enables the measurement of effectiveness, satisfaction and change in the same.

For this group, the vision for student evaluation is: 'tell me that something needs to change; enable me to measure this change'. For this group, 'success' lies in seeing the metrics change; a successful MEQ provides these change data.

However, the majority of participants in a teaching role expressed the view that improvement is driven by information about what needs to change. Student evaluations must feedforward, rather than feedback, to enable teaching that better meets the learning needs of the cohort giving the evaluation. This consensus on the purpose of SET, across teaching staff was observed across all disciplines and Levels.

For this group, the vision for student evaluation is: 'tell me what needs to change'. For this group, a successful MEQ does not (just) measure change. Rather, 'success' is defined as having an increased understanding of student needs and having the ability to make evidence-based changes to practice, going forward.

These findings closely parallel Matthews et al (2018), a study which contrasts institutional senior leaders' perceptions of the meaning of students as partners with those of academics and students. The authors find that senior leaders engage with the idea of the student voice through the lens of metrics that capture student satisfaction for quality assurance, rather than capturing feedback for enhancement; as one-way, rather than dialogue; and as important for institutional reputation and funding outcomes.

Stage 2 failed to develop a consensus. As a result, three questionnaires were developed and piloted in Stage 3. This aimed to allow an evidence-based decision on which questionnaire would gather the most useful, usable data, for both academics and senior leaders.

Stage 3: Questionnaire pilot

Participants were randomly assigned to complete one of the three questionnaires. The first questionnaire was purely quantitative, consisting of 11 questions, which students could rate on a 1-5 Likert scale, plus don't know/not applicable. The questions were written in an objective style, asking students to rate aspects of their taught experience. Broadly, this questionnaire appealed more to senior leaders.

Questionnaires 2 and 3 appealed, broadly, more to academics. Questionnaire 2 was quantitative, consisting of 13 questions, written in a more subjective style, asking students how they felt about their learning experience. Questionnaire 3 contained 3 questions, each allowing an open, free-text response, asking simply what worked well for the student, what worked less well for the student and what one thing the student would change, if they could. This questionnaire was initially designed as part of questionnaire 2, but was separated out, because of concerns over inverse relationship between questionnaire length and response rates (Nicolaou and Atkinson, 2019: 98).

Stage 3 recruited a maximum variation sample, considering Faculty, Level, discipline, module size, session type and campus. Faculty and Level are given in Table 1. The sample included 19 distinct disciplines and a range of module sizes, from 2 to 200, with a mean average module size of 34. Session types include lectures, workshops, seminars and laboratories, across two campuses. Recruitment of students was via module tutors. Tutors were initially informed about the study through a general call for participants, then via a more targeted, direct approach from the study author, to fulfil sample needs for maximum variation. The study author explained the study to students during a taught session and all students registered on the module were sent a follow-up recruitment email from the author, via their tutor.

Questionnaire	Students (n)	Modules (n)	Level* (n)					Faculty (n)			
			0	4	5	6	7	Arts / Hums	Education	Health / Wellbeing	Social and Applied Sciences
1	45	10	1	3	2	2	2	2	4	2	2
2	43	13	2	4	3	2	3	2	3	2	6
3	48	12	1	4	3	1	3	2	4	2	4
Total	136	35	4	11	8	5	8	6	11	6	12

*N=36 because one module was taught at Level 6 and 7.

Table 1. Stage 3: student sample.

The utility of the questionnaires is considered in Stage 4, below. In this section, we focus on the student voice in the SET design.

The literature review revealed that students perceive a different purpose of partnership, voice and SET. Consequently, students have different needs from an MEQ. Therefore, after completing the MEQ, all participants were asked the following, additional questions:

- Thinking about the questions you've just answered, did they let you say what you wanted to say about your module? Yes/no.
 - If no, what else would you have liked to tell us?
- Do you think that the questions asked were the right ones to help us make positive improvements to your module? Yes/no.
 - If no, why not – and what questions would you like module evaluation questionnaires to ask, which could make a difference?

The majority of students who commented on the pilot questionnaires expressed a desire to discuss their learning experience. There was a strong desire to explain the reasons for quantitative ratings, to explain what worked, what did not work, the reasons why and, crucially, what would work better. The feedback suggests that students value being able to express a direct, anonymous and confidential voice, allowing them to give honest, unstructured feedback on the aspects of their learning experience that matter to them. These participants wanted to be able to comment freely on what mattered to them, not to be restricted to a list of pre-determined topics, even when this list included the topic of concern.

The order of the questions emerged as important. This research finds that students feel that their opinion is valued more through qualitative questions than through quantitative questions. Placing qualitative questions before quantitative questions enhances student perceptions of the value of their voice.

For this group, the vision for student evaluation is: 'work with me to change'. For this group, a successful MEQ allows the unfettered student voice, builds relationships through the suggestion of a partnership for enhancement and allows a thoughtful expression of the student's learning experience. It empowers the student voice, improving students' ability to think clearly and critically about the learning process, enabling students to be true partners in learning.

Stage 4: Module leader consultation

In Stage 4, we presented the results from all three questionnaires to academics and senior leaders. They were asked to consider what action they would take, if they had received each of the three set of results. At the end of the discussion, participants were asked to select which combination of questions would best enable learning and teaching enhancement and student partnership in learning.

To capture the academic voice, the leaders of the 36 modules identified in Stage 3 were contacted. 19 replied. 11 sent comments by email; 8 were interviewed. The participant profile is given in Table 2.

In addition, 9 senior leaders of learning and teaching from across the University participated in a discussion group. As above, they were asked to review the results, consider what action they would take if they received these results and select the combination of questions that would best enable

learning and teaching enhancement. Participants were from all four Faculties, Central Services and Professional Services.

	Level					Faculty				Questionnaire		
	0	4	5	6	7	A&H	Ed	HWB	SAS	1	2	3
Number of participants	3	6	3	4	5	4	6	3	8	7	9	4

Table 2. Stage 4: Module Leader sample.

After considering the results, module leaders were unanimous in their demand for open, qualitative questions. It emerged strongly that, for this group, only qualitative data can provide the detail needed for module enhancement. Furthermore, reflecting findings reported in Yiu et al (2019), module leaders valued these data because they provided specific, evidence-based feedback, close to the point of delivery, which highlighted areas for professional development.

Students' qualitative responses focused overwhelmingly on perceptions of teaching methods, teaching resources, teaching staff and subject content. The study revealed unanticipated areas of student feedback, which no module leader had foreseen. These data suggested that there are minor enhancements that could be made, which could make substantial difference to satisfaction, but which could not have been revealed through closed, quantitative questions, either because module leader would not have thought to ask these questions, or because including these questions would result in a lengthy, unwieldy questionnaire. Furthermore, the wide range of themes that emerged from the qualitative Questionnaire 3, which cannot be mapped to the quantitative questions asked in Questionnaires 1 and 2, suggests that students have much more to contribute to module improvements than our current MEQs allow.

Module leaders did not find the volume of qualitative data to be overwhelming. All expressed that the qualitative data forced a deeper engagement with feedback than would occur with quantitative data. To draw an analogy with student interactions with feedback, our colleagues were unable to see their 'grade' without first interacting with their feedback.

This said, notwithstanding the strong preference for qualitative data, many module leaders also requested quantitative data. After reviewing the evidence, it emerged naturally and strongly that qualitative data alone is insufficient to inform enhancement. Quantitative data would give perspective regarding the scale of satisfaction/dissatisfaction, plus give the ability to track the impact of improvements.

After reviewing the evidence, senior leaders remained committed to the need for SET to gather quantitative performance monitoring data, to enable change to be measured and accountability to be enforced. However, through discussion, it emerged strongly that such data could not guide enhancement. Whilst quantitative metrics can indicate a need for improvement, they cannot inform actions. Participants were only able to design interventions after reviewing the qualitative data. As such, participants took the evidence-based decision to gather qualitative data, which will provide evidence to support the enhancement of teaching excellence, which will in turn feed into improvements in scores recorded in national benchmarking surveys.

Thus, we reached a compromise, which was guided and facilitated by the student voice. In the proposed MEQ, qualitative data will give valuable insights into how to enhance teaching excellence.

Quantitative data allows progress to be tracked. And students can provide an unfettered voice, overcoming concerns expressed in Zepke (2018) that SETs exist to reinforce the status quo and providing qualitative evidence for enhancement, whilst simultaneously providing quantitative data on key metrics to enable monitoring and a degree of accountability.

The proposed Module Evaluation Questionnaire

The proposed MEQ is given in Figure 2, below.

1.	What worked well on the module?					
1.	What worked less well?					
1.	If you could change one thing about the module to make it better for you, what would it be?					
		Strongly disagree	Disagree	Agree	Strongly agree	Don't know / Not applicable
1.	Overall, I was satisfied with this module.					
1.	I understood the aims of this module.					
1.	I was given useful feedback on how to improve, throughout the module.					
1.	I felt involved and engaged in my learning.					

Figure 2. Module Evaluation Questionnaire for effective monitoring, enhancement and relationships.

The proposed MEQ includes a small number of quantitative questions, with the option to add up to two further quantitative questions of the module leader's own design. Adding a small number of quantitative questions addresses the data gaps that emerge from a purely qualitative survey, ensuring that we gather feedback from students on all important areas of L&T. The quantitative questions that have been chosen focus on four areas that are key to enhancement, but which did not naturally emerge in qualitative data. These are: overall satisfaction; module aims / learning outcomes; feedback; and engagement.

The proposed MEQ does not include questions included within other Institutional surveys. It does not include questions from National (UK) surveys, such as the National Student Survey (NSS)¹. Neither staff nor students believe that such questions will provide the evidence needed to improve modules and enhance teaching excellence, nor do they believe that it is appropriate to use MEQs to practise for national surveys. Evidence suggests that to do so would reduce response rates and would reduce the quality of the qualitative data collected. However, using the findings from Stage 3 of this research, senior leaders can map these responses to the majority of NSS questions, to

¹ <https://www.officeforstudents.org.uk/publications/the-national-student-survey-2019/>, viewed 23 October 2018.

demonstrate that an uplift in NSS scores is highly likely, if action is taken to address concerns revealed by the qualitative questionnaire.

Conclusion

This research validates the claim that student evaluations of teaching have great potential to facilitate the student voice, creating a genuine student-staff partnership to enhance what and how we teach and, therefore, what and how our students learn.

Results suggest that, in the main, different actors in HE have different visions of the purpose of SET. Whilst all actors focus on improvement, the perception of what improvement looks like, alongside perceptions of the way to drive improvement, differ. Therefore, perceived data needs to facilitate improvement also differ. They therefore have different demands from MEQs.

These, broadly, fall into three categories.

- (1) Monitoring.
- (2) Enhancement.
- (3) Relationships.

It is essential that we see all of these views as valid, not only in the interests of progressing the policy debate, but also with regard to academic rigour. As illustrated in the review of the research literature regarding the benefits, disbenefits and purposes of SET, above, there is a lack of consensus to justify the privileging of one view over another. As such, the proposed MEQ was designed to balance and meet all of these demands, developing a tool for SET that enables monitoring, enhancement and relationships. By allowing students to provide an unfettered voice, the MEQ provides qualitative evidence for enhancement, whilst simultaneously providing quantitative data on key metrics to enable monitoring and a degree of accountability.

Some aspects of this study may limit the generalisability of the results. Conducted at a single HEI in the UK, the reader may wish to consider the applicability of these findings to their own setting. The sample was, to a degree, self-selected and, by design, was not random. As a result, the Studies 2-4 may have attracted those with a higher-than-average interest in the topic. The Studies were conducted sequentially. As a result, it is possible that participants' attitudes were softened or otherwise influenced by repeated exposure to discussion and debate; and that reactions to / perceptions of the utility of the MEQ or the results will have been influenced by this. Users outside of this study will not be influenced in this way.

The influence of these factors can only be known after a large-scale roll-out. Whilst emerging results from a small-scale technical roll-out are encouraging, with anecdotal evidence suggesting increased response rates and greater utility of responses, perhaps reflecting the increase in salience of the MEQ as perceived by staff and students (Thielsch et al, 2018), this must be treated with caution. Further testing of the proposed MEQ, in large-scale roll-out, is the next step in this research.

References

- Alhija, F. N. A. 2017. Guest editor introduction to the special issue 'Contemporary evaluation of teaching: Challenges and promises'. *Studies in Educational Evaluation*. 54. 1-3.
- Alhija, F. N. A. 2017b. Teaching in higher education: good teaching through students' lens. *Studies in Educational Evaluation*. 54. 4-12
- Biggs, J. and Tang, C. 2011. *Teaching for quality learning and university*. Fourth Edition. Buckinghamshire: Open University Press.
- Boring, A. 2017. Gender biases in student evaluations of teaching. *Journal of Public Economics*. 145. 27-41.
- Chan, R. Y. 2016. Understanding the purpose of higher education: an analysis of the economic and social benefits for completing a college degree. *Journal of Education Policy, Planning and Administration*. 6: 5. 1-40.
- Creswell, J. W. and Plano Clark, V. L. 2007. *Designing and conducting mixed methods research*. London: Sage Publications Ltd.
- Darwin, S. 2017. What contemporary work are student ratings actually doing in higher education? *Studies in Educational Evaluation*. 54: 13-21.
- Healy, M. Flint, A. and Harrington, K. 2014. Engagement through partnership: students as partners in learning and teaching in higher education. Available online via: https://www.heacademy.ac.uk/system/files/resources/engagement_through_partnership.pdf, viewed 24 April 2019.
- Higher Education Funding Council for England (HEFCE). 2017. The well-being of graduates: assessing the contribution of higher education to graduates' wellbeing in the UK. Available online via: <https://webarchive.nationalarchives.gov.uk/20180319114412/http://www.hefce.ac.uk/pubs/year/2017/201731/>, viewed 18 April 2019.
- Kahu, E. R. 2013. Framing student engagement in higher education. *Studies in Higher Education*. 38:5. 758-773.
- Kenyon, S. 2015. Development and social policy: the role of transport in the UK context. In Hickman, R., Bonilla, D., Givoni, M. And Banister, D. Eds. *International Handbook on Transport and Development*. Cheltenham: Edward Elgar. 430-440.
- Linse, A. R. 2017. Interpreting and using student ratings data: guidance for faculty serving as administrators and on evaluation committees. *Studies in Educational Evaluation*. 54. 94-106.
- Liu, YU., Visone, J., Mongillo, M. B., and Lisi, P. 2019. What matters to teachers if evaluation is meant to help them improve? *Studies in Educational Evaluation*. 61. 41-54.
- Lord, C. G., Ross, L. and Leper, M. R. 1979. Biased assimilation and attitude polarization: the effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*. 37:11. 2098-2109.
- MacNell, L., Driscoll, A. and Hunt, A.N. 2015. What's in a name: exposing gender bias in student ratings of teaching. *Innovative Higher Education*. 40:4. 291-303.

- Matthews, K. E., Dwyer, A., Russell, S. and Enright, E. 2018. It's a complicated thing: leaders' conceptions of students as partners in the neo-liberal university. *Studies in Higher Education*. Available online via: <https://www.tandfonline.com/doi/abs/10.1080/03075079.2018.1482268>, viewed 24 April 2019.
- Mitchell, M. W. and Martin, J. 2018. Gender bias in student evaluations. *PS: Political Science and Politics*. 51:3. 648-652.
- Nicol, D. J. and Macfarlane-Dick, D. 2006. Formative assessment and self-regulated learning: a model and seven principles of good feedback. *Studies in Higher Education*. 31:2. 199-218.
- Nicolaou, M. and Atkinson, M. 2019. Do student and survey characteristics affect the quality of UK undergraduate medical education course evaluation? A systematic review of the literature. *Studies in Educational Evaluation*. 62. 92-103.
- Nielsen, T. and Kreiner, S. 2017. Course evaluation for the purpose of development: what can learning styles contribute? *Studies in Educational Evaluation*. 54. 58-70.
- Office for Students (OfS). 2018. The Teaching Excellence and Student Outcomes Framework (TEF): a short guide to the 2018 awards. Available online via: https://www.officeforstudents.org.uk/media/5ff81204-14f4-4e71-8b48-91f46247c49b/tef_short_guide_2018.pdf, viewed 18 April 2019.
- Office for Students (OfS). 2018b. The Teaching Excellence and Student Outcomes Framework: subject level pilot guide. Available online via: https://www.officeforstudents.org.uk/media/57eb9beb-4e91-497b-860b-2fd2f39ae4ba/ofs2018_44_updated.pdf, viewed 25 April 2019.
- Office for Students (OfS). 2019. The Teaching Excellence and Student Outcomes Framework (TEF): the TEF year 4 metrics. Unpublished.
- Spooren, P. and Christiaens, W. 2017. I liked your course because I believe in (the power of) student evaluations of teaching (SET). Students' perceptions of a teaching evaluation process and their relationships with SET scores. *Studies in Educational Evaluation*. 54. 43-49.
- Thielsch, M., Brinkmoller, B. and Forthmann, B. 2018. Reasons for responding in student evaluation of teaching. *Studies in Educational Evaluation*. 56. 189-196.
- Utriainen, J., Tynjala, P., Kallio, E. and Marttunen, M. 2018. Validation of a modified version of the Experiences of Teaching and Learning Questionnaire. *Studies in Educational Evaluation*. 56. 133-143.
- Uttl, B., White, C. A. and Gonzalez, D. W. 2017. Meta-analysis of faculty's teaching effectiveness: student evaluation of teacher ratings and student learning are not related. *Studies in Educational Evaluation*. 54. 22-42.
- Zepke, N. 2015. Student engagement and neo-liberalism: mapping an elective affinity. *International Journal of Lifelong Education*. 34: 6. 696-709.
- Zepke, N. 2018. Student engagement in neo-liberal times: what is missing? *Higher Education Research and Development*. 37:2. 433-446.