STUDENT SURVEYS OF TEACHING & LEARNING QUALITY

An issues paper to contextualise current practices in higher education

January 2017
1 Background

The practice of surveying students about their perceptions of teaching, learning, course (unit), and program (award/degree) quality is well-established within Australian and international universities (Alderman, Towers, & Bannah, 2012; Barrie, Ginns, & Symons, 2008; Hirschberg, Lye, Davies, & Johnston, 2011; Tucker, 2013). Within Australia, student evaluations of teaching conducted via a national instrument, the Course Experience Questionnaire (CEQ), was originally developed to function as an indicator of universities’ teaching performance (Linke & Performance Indicators Research Group, 1991). Ramsden (1991a) stated that “Student ratings of teaching measure a key attribute of higher education institutions and their component units - the quality of their instructional practice, curriculum, and assessment” (p.1). Earlier studies by Marsh (1987), Cashin (1988), and Entwistle and Tait (1990), and subsequent studies by Feldman (2007), Marsh (2007), Barrie et al. (2008), Hirschberg et al. (2011), among many others, have similarly concluded that student evaluations help to inform improvements to the quality of teaching practices, and curriculum content and activities. The subsequent development of institution-based survey instruments across all universities within Australia further facilitated student evaluations of teaching to provide staff with evidence to support re-employment, tenure, promotions, awards, and performance management and development (Barrie et al., 2008).

At the national level, surveys are conducted annually by the Australian Government to address three purposes:

1. To ensure adequate information is available for students to make informed decisions about their study options.
2. Higher education providers will receive data from the surveys relating to their students and graduates, supporting their continual improvement efforts in key areas such as teaching practices, learner engagement and student support.
3. The indicator suite will provide information on the student experience and employment prospects, enabling Australia to benchmark performance against the United States of America, United Kingdom and New Zealand. (Department of Education and Training, 2015)

Similar to the European Commission’s U-Multirank1, these surveys are part of the Quality Indicators for Learning and Teaching (QILT) initiative2 which aim to promote the quality of Australian higher education institutions to local as well as international students. These surveys act as a measure of quality to promote the prestige of Australian universities by focusing “on quality teaching and [enabling] Australia to internationally benchmark performance” (Department of Education and Training, 2015).

At the institutional level, practices, methodologies, and methods can vary considerably as they are contextualised to the institution’s needs (Barrie et al., 2008; Hirschberg et al., 2011; Nulty, 2008). Recent changes made by some universities to their student survey systems have been designed to move institutional dependencies on single data sources to a more holistic approach to demonstrate teaching and learning quality (see: Alderman, Bennett, & Phan, 2013; FASS Teaching

1 http://www.umultirank.org/
2 https://www.qilt.edu.au/
and Learning, 2015; Institute for Teaching and Learning, 2015). Support for a broader evidence base for teaching and learning quality has also been provided nationally by the Office for Learning and Teaching (OLT) through the funding of the Australian University Teaching Criteria and Standards Project. Internationally, the UK’s Higher Education Academy (HEA) has also supported the move to a broader evidence base through the development of the Promoting Teaching: Evidence Framework (HEA, 2013). In order to help contextualise these current practices in higher education, this paper provides a summary of some key issues surrounding the usage of student surveys of teaching and learning quality.

2 Impacts on Teaching Quality

The premise on which student evaluations of teaching is founded asserts that the student “voice” (that is, the student perception) is a valuable data source for evidencing teaching impacts on student learning (Cashin, 1988; Dawkins, 1988; Entwistle & Tait, 1990; Marsh, 1987; Ramsden, 1991b). This is still widely accepted within the literature reviewed with many authors noting that thousands of studies have been conducted both within Australia and internationally (see: Alderman et al., 2012; Zumrawi, Bates, & Schroeder, 2014). Student evaluations have been shown to have direct impacts on teaching quality by affecting changes which help to improve teaching practices; helping staff to refine the way they teach and how they teach (Barrie et al., 2008; Biggs & Tang, 2011; Feldman, 2007; Kember & Gins, 2012; Smith, 2008). Student evaluations of teaching have been further shown to improve the relationship or rapport between teachers and students by identifying whether information was conveyed clearly to student, or whether staff had taught with enthusiasm or were approachable (Balam & Shannon, 2009; Biggs & Tang, 2011; Hirschberg et al., 2011; Kember & Gins, 2012; Smith, 2008). Refinements to survey instruments across Australian universities have further helped student evaluations of teaching to impact a staff member’s teaching practice by surfacing perceived best practice and areas for improvements (Hirschberg et al., 2011). As a consequence, student satisfaction with individual staff member’s teaching has continued to function as an evidence base for staff to support their applications for teaching awards and citations at the institution level, nationally, and internationally (see: Office for Learning and Teaching, 2014; HEA, 2015).

Although the benefits of student evaluations of teaching have been continually shown to directly impact teaching quality, there are risks or other negative effects which have been highlighted by some authors, such as Kember, Leung, and Kwan (2002), Richardson (2005, 2009), Davies, Hirschberg, Lye, Johnston, and McDonald (2007), and Ongeri (2009). Risks or negative effects outlined have included the inverse effects to the positive impacts on teaching quality. These inverse effects included continuing declines in teaching performance and a break-down in the relationship between students and staff (Davies et al., 2007; Kember et al., 2002). These effects are exacerbated when institutions do not implement robust policies to support staff nor offer counselling services when needed to deal with negative student feedback (Kember et al., 2002).

3 Impacts on Course Quality

Student evaluations of course quality have also been key components of national surveys, such as the CEQ (Hirschberg et al., 2011; Ramsden, 1991a). Refinements to the Graduate Destination Survey (GDS) were further designed to assist institutions to evaluate the quality of programs from a whole-of-program approach (Graduate Careers Australia, 2013). Institutional practices
throughout Australia have embedded evaluations of courses into course quality assurance mechanisms (Alderman et al., 2012; Hirschberg et al., 2011; Ramsden, 1991b). These instruments have sought to collect feedback from students about specific components within courses, such as assessment, curriculum, the sufficiency of feedback to students, or learning support (Carroll, 2013; Hirschberg et al., 2011; Kember & Ginns, 2012). They have been further shown to affect changes in a positive manner (Barrie et al., 2008; Hirschberg et al., 2011; Ramsden, 1991b) with flow-on ramifications to student learning (Barrie et al., 2008; Cathcart, Greer, & Neale, 2013).

However, like the student evaluations of teaching, course evaluations have also been shown to impact negatively on courses. Arguments about the negative impacts include risks related to students not knowing what they don’t know. This is held in the belief that students’ perceptions are not akin to that of content or discipline experts (teachers) (Biggs & Tang, 2011; Kember et al., 2002; Porter, 2011). For these reasons, many universities within Australia, as well as international bodies, are now advocating a holistic approach to evidencing course quality (Alderman et al., 2013; FASS Teaching and Learning, 2015; Institute for Teaching and Learning, 2015; HEA, 2013).

4 Key Issues

A scan of literature (published from 2010 onwards) addressing student surveys using search terms, such as surveys, feedback, evaluation, and questionnaire found that the key issues listed below frequently appeared within the various database searches (Google Scholar https://scholar.google.com.au/ and Taylor & Francis Online http://www.tandfonline.com/). Key issues based on the aggregated citations of the listed search terms were:

- Validity and reliability
- Response rates
- Paper vs online
- Incentives
- Impact on Student Learning

Although these keys issues will be addressed through separate sub-sections below, some are interrelated issues and directly impact others.

An additional section on the type of survey (measures of learning vs. perceptions of quality) has also been included to discuss recent meta-analysis findings in regard to teaching evaluations.

4.1 Validity and Reliability

The majority of the literature reviewed focused on the validity and reliability of student surveys of teaching. The key findings from these studies generally concluded that student evaluations of teaching are a valid and reliable measure of teaching quality when certain survey conditions or outcomes have been met. These include questions that are clear and provide students with the ability to leave written feedback (Porter, 2011; Skowronek, Friesen, & Masonjones, 2011;Spooren, Brockx, & Mortelmans, 2013), and the achievement of a proportionate response rate, including the usage of data aggregation across courses/units for surveys related to a teacher (James, Schraw, & Kuch, 2014; McCullough & Radson, 2011; Zumrawi, Bates, & Schroeder, 2014).

These studies further warned of decreased validity and reliability when survey questions were: too broad or vague (Ernst, 2014; Porter, 2011; Spooren et al., 2013); were used in isolation from other forms of evidence (James et al., 2014; Spooren et al., 2013); were over-interpreted,
misinterpreted, or misused (Ernst, 2014; James et al., 2014; McCullough & Radson, 2011); or, had achieved poor response rates (Ernst, 2014; James et al., 2014; Rantanen, 2012; Spooren et al., 2013; Zhao & Gallant, 2011; Zumrawi et al., 2014).

4.2 Response Rates

In their recently published paper Zumrawi et al. (2014) found that the desired response rates for class sizes ranging from 10 to 1,000 students gradually decreased from just under 100% for classes with ten students to just under 10% for classes with 1,000 students (using a 5% margin of error, 90% confidence level, and with P= 0.9 where P is the probability of a favourable response). These results were modelled on four years of student evaluation data from the University of British Columbia, Canada, which achieved an average of over 160,000 responses per survey item over the four years.

If institutions adopt an enterprise-wide response rate target, this variability in range for response rates according to class size supports the need to adopt a target which has been contextualised to the institution’s class sizes. The adoption of institution-wide response rate targets, whether formally or informally, is exemplified in Australia by the University of Melbourne. It adopted a response rate target of 50% for 2014. By the first half of 2014, the University of Melbourne had achieved a response rate of 47.4%. It appears that a new cash incentive competition run in 2014 may have had a positive effect on their response rate (Teaching and Learning Quality Assurance Committee (TALQAC), 2014).

Bennett and Nair (2011), Crews and Curtis (2010), and Williams (2014) believe that it is important to understand the reasons for low response rates. The primary reason provided by Bennett and Nair was that “low response rates are highly correlated to a lack of follow-up action” (2011, p. 123). If students perceive that there is no value in completing a survey because they do not see evidence of how it impacts them, then they are less likely to complete the survey (Crews & Curtis, 2010; Harvey, 2011). To this end, some universities use a mid-semester, formative evaluation tool that aims to affect changes which are applicable to students while they are still enrolled in a course/unit. Examples of these surveys include The Australian National University’s (ANU) mid-semester feedback forms (Planning & Performance Measurement Division, 2013) and Queensland University of Technology’s (QUT) Pulse survey (Alderman, Bennett, & Phan, 2014).

4.3 Paper vs Online

If actions taken by universities to address feedback collected from students can increase response rates, then the evidence to support the move to online surveys also increases since the affordances of technology includes ease of access to timely reports, analytics, and detailed statistical analyses for qualitative and quantitative responses (Crews & Curtis, 2010; Morrison, 2013; Risquez, Vaughan, & Murphy, 2014). As more institutions worldwide switch from paper to online student evaluations, recent research by Rienties (2014) found that there was still scepticism among academic staff at the University of Surrey, UK, about the benefits of online student evaluations. This scepticism existed even though: the University of Surrey had achieved similar response rates at 60% after their transition from paper to online; and, the new online survey had yielded twice as much written feedback from students. The fear of decreased response rates when switching from paper to online was also quashed by Kordts-Freudinger and Geithner (2013) in their study which found that the evaluation situation (in class versus after class) had more impact on survey results than the evaluation mode (paper versus online).
With the increasing agility of analytical, statistical, and reporting software, the cost-benefits of switching from paper to online are also being outweighed by the swiftness of timely feedback from students and back to students. This increased efficiency is evident at QUT which releases outcomes of student surveys to staff 48 hours after closure (Alderman et al., 2013). Overwhelmingly, the recent literature is also confirming that online surveys provide students with a greater ability to provide extended written comments which are more easily analysed by staff than handwritten comments (Crews & Curtis, 2010; Morrison, 2013; Palmer & Smith, 2012; Rienties, 2014; Risquez et al., 2014).

4.4 Incentives

The desire to increase response rates has led to the adoption of incentives designed to entice students to complete their surveys. However, some authors (Bennett & Nair, 2011; TALQAC, 2014) agree that “[t]he evidence on incentives has been somewhat varied, with no indication that an incentive always results in a higher response rate” (Dommeyer et al., 2004 in Bennett & Nair, 2011).

This increasing adoption of incentives is evident in the practices of many Australian universities, including The University of Melbourne³ which gives $500 cash prizes, and QUT⁴ which donates 10 cents per survey submitted to fund student support programs. Practices in the United States extend to giving students bonus points for their assessments or other course/unit based benefits, such as early release of grades (Crews & Curtis, 2010; Goodman, Anson, & Belcheir, 2014). However, suggested strategies to improve response rates extend beyond incentives. Specifically, academic staff engagement in the processes for reminding students to complete their surveys also yield better response rates (Crews & Curtis, 2010).

4.5 Impact on Student Learning

Interest in the impact on student learning as a result of engagement with surveys is also growing. This interest seems to correlate with the affordances of technology in providing ease of access to information about students, their engagement with surveys, and their perceptions of the quality of teaching and learning while still enrolled in the course/unit (Cathcart, Greer, & Neale, 2013; Tucker, 2013). A correlation can be established between the changes to evidencing impact on student learning, to student evaluations, and to data analytics. Siemens, Dawson, and Lynch (2013) argued that “Education needs new ways of thinking, new ways of doing and new ways to evaluate and demonstrate impact. The coupling of new models of pedagogy with learning analytics offers much value to the contemporary education sector” (p. 5). This is also shown within QUT’s evaluation framework which endorses different methods for collecting student achievement, engagement, and evaluation data (Alderman et al., 2013) to inform academic priorities for teaching and learning evaluations of quality.

Recommendations for the usage of student surveys are progressively moving towards including questions about student learning in order for academic staff and institutions to implement changes which directly impact the students who provided feedback (Alderman et al., 2013; Cathcart et al., 2013; FASS Teaching and Learning, 2015; Tucker, 2013). Examples of student-focused questions include:

I am taking advantage of opportunities to learn in this unit (Alderman et al., 2013, p. 8).

I make best use of the learning experiences in this unit (Curtin University, 2015).

The recent changes to the survey instrument at the University of Sydney further show the University’s move to the adoption of practices which prioritise a broader evidence base to support improvements to teaching and learning quality (FASS Teaching and Learning, 2015; Institute for Teaching and Learning, 2015). This includes changes to the “lens” (also known as “voice”) from which staff can view perceptions of quality as they impact the student and the teacher. These options, adopted from Brookfield (1995), include the Self lens, Student lens, Peer lens, and Literature lens (FASS Teaching and Learning, 2015). The University of Sydney also endorses the usage of customised survey instruments selected to match teaching modes and needs, such as surveys designed specifically for teaching in lectures, tutorials, clinics, demonstrations, or online (Institute for Teaching and Learning, 2015a).

4.6 Measures of Learning vs. Perception of Quality

A recent meta-analysis by Uttl, White and Gonzalez (2016) disputed previous findings (Cohen 1980, 1981; Feldman, 1989) of the moderate correlation between Student Evaluation on Teaching (SET) ratings and learning. The meta-analysis re-analysed previously published meta-analyses and concluded that in large sample sizes there was no significant correlation between Student Evaluation on Teaching (SET) ratings and learning, and thus students do not learn more from highly rated teachers.

Although this finding may have potential implications for surveys that are evaluating measures of learning, it must be noted that many survey instruments at Australian universities are now only measuring students’ satisfaction or perceptions of course or teaching quality5. At The University of Queensland, the Student Evaluation of Course and Teaching survey (SECaT) is not considered a ‘measure of learning’ instrument. The SECaT survey primarily measures students’ perception of course and teaching quality, rather than their perceptions of how much they have learned. In 2016, The SECaT Working Party further recommended that one item that could be considered a ‘measure of learning’ construct (Q7. I learned a lot in this course), be removed from the instrument to ensure the SECaT remains focused as a perception of quality instrument.

6 References


