University scholarship activities on module evaluation questionnaires at a British University

Junko Winch
University of Sussex, Brighton, UK

Abstract

Purpose – The purpose of this study comprises the following three: (1) to ascertain the purpose of university module evaluation questionnaires (MEQs) and its reliability; (2) to evaluate University X's MEQ; and (3) to offer how Universities may be able to support their teaching staff with scholarship activities using the MEQ project.

Design/methodology/approach – University MEQ purposes and its reliability were investigated using literature reviews. The University X's MEQ seven statements were evaluated by three university academic staff. The study was conducted at a British university in South East of England. The duration of this interdisciplinary project was for two months which was a university interdisciplinary project between 14/07/20 and 13/10/20.

Findings – The purpose for MEQs includes (1) students' satisfaction; (2) accountability for university authority and (3) teaching feedback and academic promotions for teaching staff. The evaluation of University X's MEQ indicated that MEQ questions were unclear which do not serve reliable student evaluation results. This topic may be of interest to University MEQ designers, lecturers, University Student Experience team, University Executive Board, University administrators and University HR senior management teams.

Originality/value – The following three points are considered original to this study: (1) MEQ purposes are summarised by students, university authority and teaching staff; (2) the evaluation of a British University MEQ; (3) provides suggestions on how lecturers' scholarship activities can be supported by the university-wide initiative and umbrella network. These are practical knowledge for the faculty and administrators of higher education institutions which may be of use.

Keywords Academic promotion, Assessment of teaching, Module evaluation questionnaires, Student evaluations of teaching

Paper type Research paper

1. Introduction

1.1 Issues under consideration

There has been a recent shift in the teaching staff's academic responsibilities related to promotion and academic employment in the UK Teaching staff's, especially lecturers', scholarship activities are more emphasised in addition to their teaching responsibilities. Scholarship activities are individual's professional development in teaching and learning, for example, journal publications, conference presentation, research-related activities and so on. Another shift of academic responsibilities can be observed in the recruitment in the lectures' job descriptions and their pay grade. Many teaching-focus positions were advertised in non-STEM subjects, which the teaching staff are not expected to do any scholarship activities. However, the current academic trend places more emphasis on scholarship activities even for

Disclosure statement: No potential conflict of interest was reported by the author.

non-STEM lecturers' positions. This sudden change affects individual academic staff. It affects their promotion and may bring anxiety and burden to the lecturers, who occupy the majority of the university academic community. But what is this shift pertinent to MEQs and how does it affect MEQs?

The next subsection describes the background of this University X's university-wide initiative and umbrella network.

1.2 Background of the MEQ project
The above shift in the teaching staff's academic responsibilities in the United Kingdom also affected at University X, when the emphasis started just before the Covid-19 lockdown in July 2020. However, there was University X's university-wide initiative and umbrella network called "DARE to Transform" which supported the lecturers' scholarship activities by providing scholarship opportunities for the lecturers.

1.2.1 DARE scholarship programme. In April 2020, the University X's new mentoring scheme was launched, led by an Associate Dean of Education and Student, Business School and the head of Technology Enhancement Learning. DARE Scholarship programme is "a university-wide initiative and umbrella network to enhance scholarly practice, encourage educational experimentation and foster academic enquiry. It also serves as a vehicle for the promotion and sharing of the outputs and outcomes from scholarly activities from across the University X" (University X, 2021).

The purpose of the mentoring scheme was to provide more experienced staff on the Education and Scholarship track to support new or early career colleagues. This mentoring scheme benefits both experienced staff and lectures to develop their scholarship profile in scholarship activities. The mentors were recruited from Senior Lecturers or Professors on the Education and Scholarship career pathway, and the mentees were recruited from Lectures in the Education and Scholarship career pathway from any disciplines at University X.

The researcher sent my expression of interest on 01/04/20 and was selected as the first cohort (total of seven mentees) the Scholarship Mentoring programme. Both mentors and mentees participated in the first training session called "how to be successful mentees" on 21/05/20. Six mentees requested their preference of mentees whether they wish to be mentored within the same disciplines or outside the discipline. The researcher had a "chemistry meeting" (mentoring match) with a mentor outside her discipline and a signed a contract with the HR for 12 months to have meetings on a monthly basis.

MEQ project is separate from the mentoring schemes, and the membership to DARE Scholarship programme does not automatically qualify all members to work on the MEQ project. There were two further selection processes involved.

The researcher applied for the university-scale interdisciplinary MEQ project by sending a CV. The benefit of this project included: (1) a supervised project by the Associate Dean of the Business School and; (2) an interdisciplinary collaborative project. As the researcher belongs to School of Media, Arts and Humanities, she finds this project attractive to be supervised by a supervisor from the different discipline. The researcher also finds interdisciplinary collaborative project attractive as it would give opportunities to interact with colleagues from the Social Science where the researcher's original academic background (education) was from.

The researcher received an e-mail from the Associate Dean of the Business School to invite me to work on the MEQ project on 09/07/20 and collaborated working with a Deputy Course Director-International Marketing MSc from Business School.

1.3 MEQ project procedure
This was a two-month university project in the beginning of Covid-19 pandemic, conducted between 13/07/20 and 13/10/20. On 14/07/20, the researcher had a first MEQ meeting with the

Deputy Course Director-International Marketing MSc (Business School). On 23/07/20, the second MEQ meeting was held with the Deputy Course Director-International Marketing MSc and the Associate Dean of the Business School-International Marketing MSc (Business School). The Deputy Course Director-International Marketing MSc left the project after one month and the researcher completed MEQ project alone. On 06/08/20, the third MEQ meeting was held with the Associate Dean of the Business School. On 14/09/20, the researcher submitted the MEQ report to the University Survey Group on 14/09/20. On 13/10/20, the researcher was invited to present the MEQ Report and answered questions from the Survey Group. The University Surveys Group included the Pro Vice Chancellor for Education and Students, Associate Dean of the Business School and the Deputy Pro Vice Chancellor of Student Experience (i.e. University's Executive Board Group).

1.4 Research questions (RQs)
RQs comprised the following two RQ questions:

RQ1. What is the purpose of University MEQs?

RQ2. Are University X's MEQs reliable?

RQ1 was investigated by literature reviews, and RQ2 was investigated by literature review and evaluation of MEQ seven statements.

Next section starts with the literature reviews of purpose and reliability of MEQs, which is followed by methodology, discussion and conclusion.

2. Literature review: purpose and reliability of MEQs
2.1 MEQs and SETs
Student Evaluations of Teaching (SETs) may have two purposes: institutional and local. Institutional SETs are what universities set the standard questionnaires and distribute them to all schools and departments, which is considered more traditional. The majority of SET research appears to discuss on the institutional level SETs. Institutional SETs are also often relied upon as evidence cases for promotion and are an important element of a teaching portfolio.

Local SETs, on the other hand, usually refer to what each school or department set its own questionnaires in addition to the Institutional SETs. Local SETs are implemented as a complement to the institutional SETs providing insight and feedback on specific disciplinary questions. A major shortcoming of the University MEQs suffers from low response rates and is inconsistent year on year or across the university.

University's SETs are also referred to as the Module Evaluation Questionnaires (MEQs). In the United Kingdom, MEQs are the summative surveys which run at the end of a term or level of study at University X. As it belongs to an institutional type, it is referred as "University MEQs". However, local practices have grown to address the need for student feedback and many schools and departments also run their own MEQs locally (we refer it as "local MEQs") in response to the limitations of MEQs.

MEQ stakeholders may include students, staff, politicians, alumni, funding bodies and the general public (Kearns, 1998). Among these stakeholders, students, staff and university authorities are the focus of our discussion.

2.2 Students
"Students as consumers or partners" (Bienefeld and Almqvist, 2004) and the concept of "consumer satisfaction" (Blackmore, 2009; Olivares, 2003; Titus, 2008) has brought into education. Students are considered to play an important part in evaluating tutors. Yet, some

students may not appreciate how the information is used and find MEQs a pointless exercise (Bassett *et al.*, 2015).

Bias is the source of unreliability. Eleven factors have been identified as contributors to student-related bias and the source of unreality of SETs: (1) Weather (Braga *et al.*, 2014); (2) Time of the day (Feldman, 1978); (3) Lower-track student; (4) Area of expertise (science vs. humanities) (Bavishi, Madera and Hebl, 2010); (5) Students' racial (black, Asian, white) stereotype (Bavishi and Madera, 2010); (6) Tutors' physical attractiveness (Campbell *et al.*, 2005; Gurung and Vespia, 2007; Hamermesch and Parker, 2005; Riniolo *et al.*, 2006); (7) Students' image compatibility (Dunegan and Hrivnak, 2003) or initial impressions of a tutor (Tom *et al.*, 2010); (8) Students' personality traits (Patrick, 2011); (9) Tutors' personality traits (this subsection was justified); (10) Students' prior interest; (11) Students' anxiety.

2.3 University authorities

SETs provide universities evidence of internal quality-assurance processes and institutional accountability (Johnson, 2000). Institutional accountability consists of the myriads of either tangible or intangible expectations (Kearns, 1998), but accountability appears to be considered of paramount important purpose at universities not only due to the British government emphasis on university's accountability and transparency across high education sector (2017), but also accountability links to attractive concepts such as "quality", "autonomy" and "integrity", all of which give special attention and positive connotation (Stensaker and Harvey, 2011). The universities report stakeholders regularly and openly (transparency), by documenting (visibility) on the outcome of their effective teaching, research activities and institutional management.

University authorities may often have ulterior motives in the student feedback process. They might be thinking in terms of league tables, for example, the National Student Survey (NSS). Some may approach it as a Key Performance Indicator (KPI). This then might feed into management decision-making, such as hiring, firing, promotion, pay, rewards and so on: "The problem here is that a KPI may mask the real reason for the decision, which could be something subjective" (Salmon, 2009, p. 77).

As current emphasis on SET data is on quantitative rating (Likert-scale questions), which is the easiness to collect, analyse and present the SET data (Penny, 2003), this also contributes why the majority of institutions use SETs as evidence.

2.4 Staff

Staff refers to teaching staff, specifically lecturers. SETs have two major purposes for teaching staff. The main and obvious purpose of SET is feedback of teaching quality for individual tutors (Spooren *et al.*, 2013). SETs give feedback not only to individual academic staff who engaged in teaching but also for supervisory School unit/course coordinators and Head of school.

Another important purpose is academic promotion (Becker and Watts, 1999) and professional career progression as a measure of quality monitoring, administrative policymaking (Penny and Coe, 2004). There is a link between SETs and tutors' promotion (Becker and Watts, 1999) and career progression, that is, input for appraisal exercises (e.g. tenure/promotion decisions). It implies a close link among SETs, academic promotion and HR department. Beran and Rokosh (2009) claim that 62% of the academic staff feel that Department Heads and Deans make proper use of SET reports. Historically, the majority of institutions have been using SETs for promotion since the 1970s when evaluations came into use for faculty personnel decisions in hiring purposes (Becker and Watts, 1999).

Two major biases related to the tutors' grading affecting reliability of SETs are validity hypothesis and grading-leniency hypothesis. Validity hypothesis is defined as "students who

have learned more in the class will receive higher grades and will naturally rate the professor more highly because of the knowledge they have gained in the course" (Patrick, 2011, p. 241). However, Abrami *et al.* (1980) argue that the effect of grading standards on students' ratings is very small and inconsistent across different rating items.

Grading-leniency hypothesis is defined that "the professor's leniency in assigning grades favourably influences student evaluation scores" (Patrick, 2011, p. 241). Simply, tutors who give higher grades also receive better evaluation (Carrell and West, 2010; Johnson, 2003; Weinberg *et al.*, 2009). Some researchers suggest that tutors "buy" good evaluations by giving high grades (Isely and Singh, 2005; Langbein, 2008; McPherson, 2006; McPherson and Todd Jewell, 2007). As career promotion is important for tutors, some tutors may strategically adjust their grades to please the students (Braga *et al.*, 2014), and others may teach the test content as students might reward tutors (Braga *et al.*, 2014).

2.5 Other variables which affect reliability of MEQs
Tutors who teach more advanced courses are rated more favourably than teaching lower-level courses (Moritsch and Suter, 1988; Cranton and Smith, 1986; Feldman, 1978; Marsh, 1980; Goldberg and Callahan, 1991). This is due to the nature of heterogeneity in class composition in lower-level courses. Heterogeneity makes it extremely difficult for tutors to tailor to the students' preferred teaching and learning experience to satisfy all. By attempting to suit the curriculum to the majority, the instructor may easily fall into the trap of mediocrity, which inevitably dissatisfies some students (Casey *et al.*, 1995; Reid *et al.*, 1981). Furthermore, Social Science and Humanities tutors are rated more favourably than STEM subjects (Cashin, 1992; Centra, 1993). Lastly, insufficient sample size (Marsh and Roche, 1997) also affects reliability.

3. Methodology
The below illustration summarises the relationship between the research questions, methods and design.
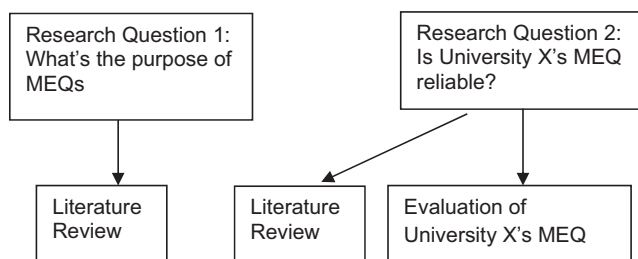
3.1 Evaluation of university MEQs
The following two points of the University MEQ seven statements were focus of evaluation: (1) whether purposes are clearly identified and (2) whether strengths and/or weaknesses of the statement were also examined in each statement. Any suggestions for changes or revisions were proposed if there are any. For the quality assurance purposes, this evaluation was carried out by three members of the MEQ project – The Deputy Course Director of International Marketing MSc (Business School), followed by the researcher and the Associate Dean of the Business School.

3.2 Literature review
The literature review involved two stages: the retrieval phase and the screening phase. In the retrieval phase, the keywords for literature review were "SETs", "reliability", "assessment" and "evaluation" using Google, Eric and ProQuest and PsycINFO. The researcher conducted all abstract and content analysis of individual article and collected initially 78 data list (60 journal papers, 17 books and 1 conference paper) from the field of psychology, education, business management and sociology (See Table A1 "Literature review list"). The dates of the journals ranged between 1978 and 2020. In the next phase, 55 qualified and chosen journals (42 journals, 12 books and 1 conference paper, which are not denoted by * in Table A1 list) were classified into the following five categories: student purpose; university authorities purpose; staff purpose; other purposes; and reliability. The selected articles and books may be biased as the researcher focussed inclusively on the specific keywords of this study for relevance. The purpose of literature review is to integrate all relevant literature related to MEQ purposes and reliability concerning students, staff and university authorities.

4. Results of MEQ seven-statement evaluation
The followings are the results of University MEQs having seven statements regarding its purpose, strength/weakness and suggestions:

4.1 MEQ statement 1: it was clear to me what I would learn and why
Purpose: To assess the clarity of the module content, but it is not clearly stated. This purpose may be suitable for all stakeholders' purposes.
  Weakness: The combination of "what and why" makes students' responses invalid.
  Suggestion: Divide the question into the following three separate questions.

I – It was clear to me what I would learn from this module.

II – It was explained to me or was stated in the module handbook why I would need to learn the content of this module.

III – The learning journey was explained to me or was stated in the module handbook why I would need to learn the content of this module.

4.2 MEQ statement 2: the materials on canvas were useful and relevant
Purpose: To assess the validity and relevance of the module content on the Virtual Learning Environment (VLE), but it is not clearly stated. This purpose may be suitable for all stakeholders' purposes.
  Weakness: One statement but includes two questions (useful and relevant). It is unclear to which content the students may answer to of whether students might find the content relevant but not useful and vice versa.
  Suggestion: Suggest to replace this to one of the followings.

I – The module content on canvas was relevant to the teaching of this module.

II – I have found the module content on canvas useful.

III – The module content on canvas helped me to learn this module better.

4.3 MEQ statement 3: the recommended reading lists were appropriate and up to date
Purpose: To assess the reading list quality, but it is not clearly stated. This purpose may be suitable for all stakeholders' purposes.

Weakness: One statement but includes two questions (appropriate and up to date). Reading lists may be too broad and vague.

Suggestion: Suggest to replace this to one of the following.

I – The reading list offered good sources that helped my learning with this module.

II – The reading list content was up to date.

III – The reading list offered a variety of resources to support my learning.

4.4 MEQ statement 4: I have received feedback on my work (including queries after teaching sessions, in person or by e-mail) which may have helped my understanding/clarified things I did not know, helped to explain a grade or identified areas for improvement
Purpose: To assess feedback. This purpose may be suitable for all stakeholders' purposes, but it is not clearly stated.

Weakness: Unclear. It asks students a mixture of multiple feedbacks in one statement. The statement refers to summative and formative feedback which can confuse students for their response. For example, a tutor may be very good with summative feedback on their module which supports students to benchmark their learning at the end of the instructional unit. However, they may fail to offer adequate formative feedback and monitor students' ongoing learning during the module.

Suggestion: Suggest replacing this to the following two-step approach questions.

The Step I question is: I have received feedback for the assessments of this module.

If answered yes to the Step I question, then one of the following three statements may be Step II question.

II-choice 1 – The feedback I received helped me to understand better, areas of improvement and my strengths.

II-choice 2 – The feedback I received explained why I was awarded with the mark on my work.

II-choice 3 – I was provided with opportunities to discuss my questions with my tutor (written and/or verbal) during the semester.

4.5 MEQ statement 5: clear information about the assessment of this module, and the marking criteria, was provided
Purpose: To assess the assessment instruction and marking criteria, but it is not clearly stated. This purpose may be suitable for all stakeholders' purposes.

Weakness: One statement but includes two questions whether information on assessment of the module was provided and if marking criteria were provided.

Suggestion: To keep the question.

4.6 MEQ statement 6: teaching accommodation and facilities were satisfactory
Purpose: To assess the learning environment, but it is not clearly stated. This purpose may be suitable for all stakeholders' purposes.

Weakness: the term "satisfaction" and "teaching accommodation and facilities" may be misleading whether it refers to learning content or physical building accommodations. This affects reliability and validity.

Suggestion: It needs confirmation regarding the meaning of "teaching accommodation and facilities".

4.7 MEQ statement 7: overall I was satisfied with the module
Purpose: To assess the students' satisfaction, but it is not clearly stated. This purpose may be suitable for all stakeholders' purposes. This affects reliability and validity.

Weakness: The term "satisfaction" may provoke different connotations and sentiment by different cultures and individuals.

Suggestion: Suggest replace this to one of the followings.

I – Overall, I feel I have learned from this module.

II – Overall, I approve the quality of teaching on this module.

III – Overall, I approve the quality of the content of this module.

IV – Overall, I approve the activities included in this module.

5. Discussion

5.1 Purpose

The purpose of seven statements was unclear. As there are only seven statements, it is assumed that the University MEQs' designers tried to include two questions in each statement, which resulted in unclear and vague questions, resulting students' various interpretations of the questions and answers. For example, questions such as "how satisfied are you . . ." do not often lead to productive answers and fail to assess teaching quality. To ensure that the intended meanings are communicated, consideration should be given to students' perception of the words such as "satisfied" as it may give different meanings depending on the students and cultures. Careful consideration should be given to the design and focus of the questions.

5.2 Reliability

Reliability is defined as "stability or consistency with which we measure something" (Robson, 2002, p. 101) and "for a research instrument to be reliable, it must be consistent" (Gray, 2004, p. 173). According to the literature reviews, MEQs remain a current yet controversial topic in higher education research and practice with many participants questioning the validity and reliability of MEQ results (Ory, 2001), and research on SETs has thus far failed to provide clear answers to the validity and reliability of SETs (Spooren *et al.*, 2013). From the evaluation of University X's MEQ statements, unclear questions imply that University MEQs as unreliable as an evaluation tool.

5.3 Strength/weakness

University MEQs allow students to provide their opinions about the course, which is one of the strengths of the University MEQs. However, the main weakness concerns the unclear purpose of university MEQ statements with vague and unclear wordings, which resulted in various students' interpretations and answers.

5.4 Suggestions

There are three suggestions. Firstly, MEQ statements should be clearly stated for students to understand the value of participating. It is suggested a staff–student partnership to agree on the purpose of the university MEQs and co-design a revised instrument, which may help to meet the stated purpose.

Secondly, students' educational cultural differences should be taken into consideration in conducting MEQs. While course evaluation is well-established practice in North America (Perry, 1990; Perry and Smart, 1997; Marsh and Dunkin, 1992; McKeachie, 1990), such practice particularly in Asia is uncommon (Marsh, 1987; Watkins, 1994). In Asia, students are

not expected to question their tutors' knowledge and authority and they are taught to respect, not criticise, their tutors (Ting, 2000).

Finally, university MEQs and academic promotion are seemingly related, relying on university MEQs as a sole and the main determinant for tutors' career progression needs to be cautious. University MEQs should be used in conjunction with other quantitative evidence such as class average attendance rate, average, maximum and minimum marks as well as qualitative response analysis which would help build a more accurate overall picture of the class and help to reduce the bias effects outlined above. Further, university MEQs and the above-mentioned teaching-related data should be provided to promotion panels to avoid the cherry picking of comments or data by applicants.

6. Conclusions

To answer RQ1, university MEQs have three purposes: (1) customer satisfaction for the students; (2) accountability for university authorities; and (3) teaching and academic promotions for teaching staff. To answer RQ2, University MEQs at University X considered unreliable. As the majority of MEQ statements are unclear in their purpose, it is possible to say that the results of MEQs are unreliable.

MEQs remain a current yet controversial topic in higher education research and practice with many participants questioning the validity and reliability of MEQ results (Ory, 2001), and research on SETs has thus far failed to provide clear answers to the validity of SETs (Spooren *et al.*, 2013). This suggests that a gap exists between the institutional reality and the literature reviews. The fact that unreliable MEQs are still heavily relied upon in academic promotion highlights issues in academic reality.

In introduction, the question was asked what the recent shift and strong emphasis on lecturers' scholarship activities affects in the introduction. The answer may be concluded that this shift may be welcoming movement to steer away from heavily relying on unreliable University MEQ results for academic promotion.

References

Abrami, P.C., Dickens, W.J., Perry, R.P. and Leventhal, L. (1980), "Do teacher standards for assigning grades affect student evaluations of instruction?", *Journal of EducationalPsychology*, Vol. 72 No. 1, pp. 107-118.

Bassett, J., Cleveland, A., Acorn, D., Nix, M. and Snyder, T. (2015), "Are they paying attention? Students' lack of motivation and attention potentially threaten the utility of course evaluations", *Assessment and Evaluation in Higher Education*, Vol. 42, pp. 431-442.

Bavish, A., Madera, J. and Bebl, M. (2010), "The effect of professor ethnicity and gender on student evaluations: judged before met", *Journal of Diversity in Higher Education*, Vol. 3 No. 4, pp. 245-256.

Becker, W.E. and Watts, M. (1999), "How departments of economics should evaluate teaching", *American Economic Review (Papers and Proceedings)*, Vol. 89, pp. 344-349.

Beran, T.N. and Rokosh, J.L. (2009), "Instructor's perspectives on the utility of student ratings of instruction", *Instructional Science*, Vol. 37, pp. 171-184, doi: 10.1007/s11251-007-9045-2.

Bienefeld, S. and Almqvist, J. (2004), "Student life and the roles of students in Europe", *European Journal of Education*, Vol. 39 No. 4, pp. 429-441.

Blackmore, J. (2009), "Academic pedagogies, quality logics and performative universities: evaluating teaching and what students want", *Studies in Higher Education*, Vol. 34, pp. 857-872, doi: 10.1080/03075070902898664.

Braga, M., Paccagnella, M. and Pellizzari, M. (2014), "Evaluating students' evaluations of professors", *Economics of Education Review*, Vol. 41, pp. 71-88.

Campbell, Gerdes, H.K. and Steiner, S. (2005), "What's looks got to do with it? Instructor appearance and student evaluations of teaching", *Journal of Policy Analysis and Management*, Vol. 24, pp. 611-620, doi: 10.1002/pam.20122.

Carrell, S.E. and West, J.E. (2010), "Does professor quality matter? Evidence from random assignment of students to professors", *Journal of Political Economy*, Vol. 118, pp. 409-432.

Casey, L., Casiello, C., Gruca-Peal, B. and Johnson, B. (1995), *Advancing Academic Achievement in the Heterogeneous Classroom*, thesis, Saint Xavier University, Massachusetts, MA.

Cashin, W.E. (1992), "Student ratings: the need for comparative data", *Instructional Evaluation and Faculty Development*, Vol. 12, p. 146.

Centra, J.A. (1993), *Reflective Faculty Evaluation: Enhancing Teaching and Determining Faculty Effectiveness*, Jossey-Bass, San Francisco.

Cranton, P. and Smith, R.A. (1986), "A new look at the effect of course characteristics on student ratings of instruction", *American Educational Research Journal*, Vol. 23 No. 1, pp. 117-128.

Dunegan, K.J. and Hrivnak, M.W. (2003), "Characteristics of mindless teaching evaluations and the moderating effects of image compatibility", *Journal of Management Education*, Vol. 27, pp. 280-303, doi: 10.1177/1052562903027003002.

Feldman, K.A. (1978), "Course characteristics and college students' ratings of their teachers and courses: what we know and what we don't", *Research in Higher Education*, Vol. 9 No. 3, pp. 199-242.

Goldberg, G. and Callahan, J. (1991), "Objectivity of student evaluations of instructors", *Journalof Education for Business*, Vol. 66, pp. 377-378.

Gray, D.E. (2004), *Doing Research in the Real World*, Sage Publications, London.

Gurung, R. and Vespia, K. (2007), "Looking good, teaching well? Linking liking, looks, and learning", *Teaching of Psychology*, Vol. 34, pp. 5-10, doi: 10.1080/00986280709336641.

Hamermesh, D.S. and Parker, A. (2005), "Beauty in the classroom: instructor's pulchritude and putative pedagogical productivity", *Economics of Education Review*, Vol. 24, pp. 369-376, doi: 10.1016/j.econedurev.2004.07.013.

Isely, P. and Singh, H. (2005), "Do higher grades lead to favorable student evaluations?", *Journal of Economic Education*, Vol. 36, pp. 29-42, doi: 10.3200/JECE.36.1.29-42.

Johnson, R. (2000), "The authority of the student evaluation questionnaire", *Teaching in Higher Education*, Vol. 5, pp. 419-434, doi: 10.1080/713699176.

Johnson, V.E. (2003), *Grade Inflation: A Crisis in College Education*, Springer-Verlag, New York, NY.

Kearns, K. (1998), "Institutional accountability in higher education: a strategic approach", *Public Productivity and Management Review*, Vol. 22 No. 2, pp. 140-156.

Langbein, L. (2008), "Management by results: student evaluation of faculty teaching and the mis-measurement of performance", *Economics of Education Review*, Vol. 27, pp. 417-428, doi: 10.1016/j.econedurev.2006.12.003.

Marsh, H.W. (1980), "The influence of student, course, and instructor characteristics in evaluations of university teaching", *American Educational Research Journal*, Vol. 17 No. 1, pp. 219-237.

Marsh, H.W. (1987), "Students' evaluations of university teaching: research findings, methodological issues, and directions for future research", *International Journal of Educational Research*, Vol. 11 No. 3, pp. 253-388.

Marsh, H.W. and Dunkin, M.J. (1992), "Student's evaluations of university teaching: a multidimensional perspective", in Smart, J. (Ed.), *Higher Education: Handbook of Theory and Research*, Agathon Press, New York, NY, pp. 139-213.

Marsh, H.W. and Roche, L.A. (1997), "Making students' evaluations of teaching effectiveness effective: the critical issues of validity, bias and utility", *American Psychologist*, Vol. 52, pp. 1187-1197, doi: 10.1037/0003-066X.52.11.1187.

McKeachie, W.J. (1990), "Research on college teaching: the historical background", *Journal of Educational Psychology*, Vol. 82 No. 2, pp. 189-200.

McPherson, M.A. (2006), "Determinants of how students evaluate tutors", *Journal of Economic Education*, Vol. 37, pp. 3-20, doi: 10.3200/JECE.37.1.3-20.

McPherson, M.A. and Todd Jewell, R. (2007), "Leveling the playing field: should student evaluation scores be adjusted?", *Social Science Quarterly*, Vol. 88, pp. 868-881, doi: 10.1111/j.1540-6237.2007. 00487.x.

Moritsch, B.G. and Suter, W.N. (1988), "Correlates of halo error in teacher evaluation", *Educational Research Quarterly*, Vol. 12, pp. 29-34.

Olivares, O.J. (2003), "A conceptual and analytic critique of student ratings of tutors in the USA with implications for tutor effectiveness and student learning", *Teaching in Higher Education*, Vol. 8, pp. 233-245, doi: 10.1080/1356251032000052465.

Ory, J.C. (2001), "Faculty thoughts and concerns about student ratings", *New Directions for Teaching and Learning*, Vol. 87, pp. 3-15, doi: 10.1002/tl.23.

Patrick, C.L. (2011), "Student evaluations of teaching: effects of the Big Five personality traits, grades and the validity hypothesis", *Assessment and Evaluation in Higher Education*, Vol. 36 No. 2, pp. 239-249.

Penny, A.R. (2003), "Changing the agenda for research into students' views about university teaching: four shortcomings of SRT research", *Teaching in Higher Education*, Vol. 8, pp. 399-411, doi: 10. 1080/13562510309396.

Penny, A.R. and Coe, R. (2004), "Effectiveness of consultation on student ratings feedback: a meta-analysis", *Review of Educational Research*, Vol. 74, pp. 215-253, doi: 10.3102/ 00346543074002215.

Perry, R.P. (1990), "Introduction to the special section", *Journal of Educational Psychology*, Vol. 82 No. 2, pp. 183-188.

Perry, R.P. and Smart, J.C. (Eds), (1997), *Effective Teaching in Higher Education: Research and Practice*, Agathon Press, New York, NY.

Reid, M.I., Clunies-Ross, L.R., Goacher, B. and Vile, C. (1981), *Mixed Ability Teaching: Problems and Possibilities*, Neer-Nelson Publishing Company, Windsor.

Riniolo, T.C., Johnson, K.C., Sherman, T.R. and Misso, J.A. (2006), "Hot or not: do professors perceived as physically attractive receive higher student evaluations?", *Journal of General Psychology*, Vol. 133, pp. 19-35, doi: 10.3200/GENP.133.1.19-35.

Robson, C. (2002), *Real World Research* –, 2nd ed., Blackwell, Oxford.

Salmon, G. (2009), *Podcasting for Learning in Universities*, Open University Press, Maidenhead.

Spooren, P., Brockx, B. and Mortelmans, D. (2013), "On the validity of student evaluation of teaching: the state of the art", *Review of Educational Research*, Vol. 83 No. 4, pp. 598-642.

Stensaker, B. and Harvey, L. (2011), *Accountability in Higher Education–Global Perspectives on Trust and Power*, Routledge, New York, NY.

Ting, K. (2000), "A multilevel perspective on student ratings of instruction: lessons from the Chinese experience", *Research in Higher Education*, Vol. 41, pp. 637-661, doi: 10.1023/A:1007075516271.

Titus, J. (2008), "Student ratings in a consumerist academy: leveraging pedagogical control and authority", *Sociological Perspectives*, Vol. 51, pp. 397-422, doi: 10.1525/sop.2008.51.2.397.

Tom, G., Tong, S.T. and Hesse, C. (2010), "Thick slice and thin slice teaching evaluations", *Social Psychology of Education*, Vol. 13, pp. 129-136, doi: 10.1007/s11218-009-9101-7.

University X (2021), DARE to Transform, [Electronic version].

Watkins, D. (1994), "Student evaluations of university teaching: a cross-cultural perspective", *Research in Higher Education*, Vol. 35 No. 2, pp. 251-266.

Weinberg, B.A., Fleisher, B.M. and Hashimoto, M. (2009), "Evaluating teaching in higher education", *Journal of Economic Education*, Vol. 40, pp. 227-261.

Further reading

Ambady, N. and Rosenthal, R. (1993), "Half a minute: predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness", *Journal of Personality and Social Psychology*, Vol. 64 No. 3, pp. 431-441.

Clayson, D.E. and Sheffet, M.J. (2006), "Personality and the student evaluation of teaching", *Journal of Marketing Education*, Vol. 28 No. 2, pp. 149-160.

Fassinger, P.A. (1995), "Understanding classroom interaction: students' and professors' contribution to students' silence", *Journal of Higher Education*, Vol. 66, pp. 82-96.

Gliner, J.A., Morgan, G.A. and Leech, N.L. (2009), *Research Methods in Applied Settings–An Integrated Approach to Design and Analysis*, Routledge, New York, NY.

Graziano, A.M. and Raulin, M.L. (2000), *Research Methods–A Process of Inquiry*, Allyn and Bacon, Boston.

Hwang, K.K. (1987), "Face and favor: the Chinese power game", *American Journal of Sociology*, Vol. 92 No. 4, pp. 944-974.

Hwang, A., Francesco, A.M. and Kessler, E. (2003), "The relationship between individualism-collectivism, face, and feedback and learning processes in Hong Kong, Singapore, and the USA", *Journal of Cross-Cultural Psychology*, Vol. 34 No. 1, pp. 72-91.

Ng, H.A. and Ang, S. (1997), "Keeping 'mum' in class—feedback seeking behaviors and cultural antecedents of Kiasu and face in an Asian learning environment", *paper presented at the Academy of Management Conference*, Boston.

Appendix

	Study	Type		Journal name
1	Abrami, P. C., Dickens, W. J., Perry, R. P., and Leventhal, L.	JA	Purpose: staff	*Journal of Educational Psychology*
2	Ambady, N., and R. Rosenthal.	JA	Purpose: students	*Journal of Personality and Social Psychology*
3	Anderson, L. W., Krathwohl, D. R.	B	Method	
4	Bassett, J., Cleveland, A., Acorn, D., Nix, M. and Snyder, T.	JA	Purpose: students	*Assessment and Evaluation in Higher Education*
5	Bavishi, A., Madera, J., and Hebl, M.	JA	Purpose: students	*Journal of Diversity in Higher Education*
6	Becker, W. E., & Watts, M.	JA	Purpose: staff	*American Economic Review*
7	Beran, T. N., and Rokosh, J. L.	JA	Purpose: staff	*Instructional Science*
8	Bienefeld, S. and J. Almqvist.	JA	Purpose: students	*European Journal of Education*
9	Blackmore, J.	JA	Purpose: students	*Studies in Higher Education*
10	Braga, M., Paccagnella, M., and Pellizzari, M.	JA	Purpose: students Purpose: staff	*Economics of Education Review*
11	Braxton, J., and Hargens, L.	B	*Purpose: others	
12	Campbell, H., Gerdes, K., and Steiner, S.	JA	Purpose: students	*Journal of Policy Analysis and Management*
13	Casey, L., Casiello, C., Gruca-Peal, B. and Johnson, B.	B	Purpose: others	
14	Cashin, W. E	JA	Purpose: others	*Instructional Evaluation and Faculty Development*
15	Cashin, W. E.	B	*Purpose: others	
16	Cashin, W. E.	JA	*Purpose: others	*Idea Paper*
17	Carrell, S. E., and West, J. E.	JA	Purpose: staff	*Journal of Political Economy*
18	Centra, J. A.	B	Purpose: others	
19	Clayson, D. E., and M. J. Sheffet	JA	Purpose: students	*Journal of Marketing Education*
20	Clegg, S., and Ross-Smith, A.	JA	*Purpose: others	*Academy of Management Learning and Education*
21	Cranton, P. and Smith, R. A.	JA	Purpose: others	*American Educational Research Journal*
22	Cunningham–Nelson, S., Laundon, M. and Cathcart, A.	JA	*Method	Assessment & evaluation in higher education
23	Dunegan, K. J., and Hrivnak, M. W.	JA	Purpose: students	*Journal of Management Education*
24	Feldman, K. A.	JA	Purpose: students Purpose: others	*Research in Higher Education*
25	Fidanza, M. A.	JA	*Purpose: others	*NACTA Journal*
26	Fassinger, P. A	JA	Purpose: students	*Journal of Higher Education*
27	Gliner, J.A., Morgan, G. A. and Leech, N. L.	B	Purpose: others	

Table A1.
Literature review list

(*continued*)

	Study	Type		Journal name
28	Goldberg, G. and J. Callahan	JA	Purpose: others	*Journal of Education for Business*
29	Graziano, A. M. and Raulin, M. L.	B	Purpose: others	
30	Gray, D. E.	B	*Reliability	
31	Gurung, R., and Vespia, K.	JA	Purpose: students	*Teaching of Psychology*
32	Hamermesch, D. S., and Parker, A.	JA	Purpose: students	*Economics of Education Review*
33	Hutchings, P., and Shulman, L.	JA	*Purpose: others	*Change*
34	Hwang, A., Francesco, A. M., and Kessler, E.	JA	*Purpose: students	*Journal of Cross-cultural psychology*
35	Hwang, K. K.	JA	Purpose: students	*American Journal of Sociology*
36	Isely, P., and Singh, H.	JA	Purpose: staff	*Journal of Economic Education*
37	Jeffrey, S. Nevid, Michael, A. Ambrose, Yea, Seul Pyun	JA	Method	Teaching of psychology
38	Johnson, R.	JA	Purpose: University	*Teaching in Higher Education*
39	Johnson, V. E.	B	Purpose: staff	
40	Kearns, K.	JA	Purpose: University	*Public Productivity & Management Review*
41	Kim, C.,Damewood, E. and Hodge, N.	JA	Purpose: students	*Journal of Management Education*
42	Langbein, L.	JA	Purpose: staff	*Economics of Education Review*
43	Lemos, M., Queirós, C.,Teixeira, and Menezes, I	JA	*Purpose: others	*Assessment & Evaluation in Higher Education*
44	Marsh, H. W.	JA	*Purpose: others	*International Journal of Educational Research.*
45	Marsh, H. W.	JA	Purpose: others	*American Educational Research Journal*
46	Marsh, H. W. and Roche, L. A.	JA	Purpose: others	*American Psychologist*
47	Marsh, H. W. and Dunkin, M. J.	B	*Purpose: others	
48	McCullough, B. D. and Radson, D.	JA	*Purpose: others	*Evaluation & Research in Education*
49	McKeachie, W. J.	JA	*Purpose: others	*Journal of Educational Psychology*
50	MacNell, L., Driscoll, A. and Hunt, A.N.	JA	*Purpose: students	*Innovative Higher Education*
51	McPherson, M. A.	JA	Purpose: staff	*Journal of Economic Education*
52	McPherson, M. A., and Todd Jewell, R.	JA	Purpose: staff	*Social Science Quarterly*
53	Moritsch, B.G. and W.N. Suter	JA	Purpose: others	*Educational Research Quarterly*
54	Ng, H. A., and Ang, S.	CP	Purpose: students	
55	Olivares, O. J.	JA	Purpose: students	*Teaching in Higher Education*
56	Ory, J. C.	JA	Reliability	*New Directions for Teaching and Learning*
57	Patrick, C. L.	JA	Purpose: students Purpose: staff	*Assessment & Evaluation in Higher Education*

(*continued*)

Table A1.

	Study	Type		Journal name
58	Penny, A. R.	JA	Purpose: University	*Teaching in Higher Education*
59	Penny, A. R., and Coe, R.	JA	Purpose: staff	*Review of Educational Research*
60	Perry, R. P.	JA	*Purpose: others	*Journal of Educational Psychology*
61	Perry, R. P. and Smart, J. C.	B	*Purpose: others	
62	Peterson DAM, Biederman LA, Andersen D, Ditonto TM, Roe K	JA	*Purpose: students	*Plos One*
63	Reid, M. I., Clunies-Ross, L. R., Goacher, B. and Vile, C.	B	Purpose: others	
64	Riniolo, T. C., Johnson, K. C., Sherman, T. R., and Misso, J. A.	JA	Purpose: students	*Journal of General Psychology*
65	Robson, C.	B	Reliability	
66	Ruane, J. M.	B	Reliability	
67	Sackstein, S.	B	*Purpose	
68	Salmon, G.	B	Purpose: University	
69	Silva, K. M., Silva, F. J., Quinn, M. A., Draper, J. N., Cover, K. R., and Munoff, A. A.	JA	*Purpose: student	*Teaching of Psychology*
70	Sonntag, M. E., Bassett, J. F., and Snyder, T.	JA	*Purpose: student	*Assessment and Evaluation in Higher Education*
71	Spooren, P., Brockx, B., and Mortelmans, D.	JA	Purpose: students Purpose: staff	*Review of Educational Research*
72	Stensaker, B. and Harvey, L.	B	Purpose: University	
73	Timmerman, T.	JA	*Purpose: students	*Journal of Education for Business*
74	Ting, K.	JA	*Purpose: others	*Research in Higher Education*
75	Titus, J.	JA	Purpose: students	*Sociological Perspectives*
76	Tom, G., Tong, S. T., and Hesse, C.	JA	Purpose: students	*Social Psychology of Education*
77	Watkins, D.	JA	*Purpose: others	*Research in Higher Education*
78	Weinberg, B. A., Fleisher, B. M., and Hashimoto, M.	JA	Purpose: staff	*Journal of Economic Education*

Note(s): Publication type: JA, journal article; B, book; CP, conference paper; *indicates paper/book that includes suitable topics but not included in this article

Corresponding author
Junko Winch can be contacted at: J.Winch@sussex.ac.uk