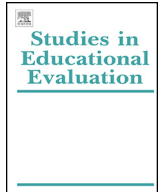




Contents lists available at ScienceDirect

## Studies in Educational Evaluation

journal homepage: [www.elsevier.com/stueduc](http://www.elsevier.com/stueduc)



# What contemporary work are student ratings actually doing in higher education?

Stephen Darwin<sup>a,b,\*</sup>

<sup>a</sup> Faculty of Education, Science, Technology and Maths, University of Canberra, Australia

<sup>b</sup> Facultad de Education, Universidad Alberto Hurtado, Chile

### ARTICLE INFO

#### Article history:

Received 15 January 2016

Received in revised form 27 May 2016

Accepted 3 August 2016

Available online xxx

#### Keywords:

Student ratings

Student evaluation

Higher education

Quantitative evaluation

Qualitative evaluation

Teacher development

Academic development

### ABSTRACT

It is conventionally assumed that student ratings perform a significant function in driving improvement in pedagogical practices in higher education. As a result, this form of evaluation has gradually become institutionalised in recent decades as an essential proxy for understanding teaching and course quality in universities across the world. However, with the rise of market-based models in higher education and heightened expectations for accountability mechanisms, the role and functional purpose of ratings-based student evaluation have become increasingly confused. This rising ambiguity has created strong tensions between the seminal drive of student ratings as a tool of quality improvement, and the emerging demands for its use as a transparent accountability measure for the comparative assessment of academic performativity. So are student ratings now largely a tool of quality assurance or performance measurement, or do they remain a legitimate tool for pedagogical improvement? This paper reports on a study that responded to this critical question by considering the contemporary work of student ratings in a major Australian university. The outcomes of this research demonstrate that tension between improvement and accountability motives is causing considerable confusion and discord around the role and value of the student voice. It also reveals that academics are tending to discount the often critical insights of students on the implications of their pedagogical practices as a result of the elevating institutional role of student ratings as a proxy for teaching quality. In considering these outcomes, rising levels of academic dissonance around student ratings would suggest a necessity to consider broadened evaluative strategies that are able to more effectively capture the improvement potential offered by the student voice.

© 2016 Elsevier Ltd. All rights reserved.

## 1. Introduction

Student evaluation—based on end-of-semester student ratings instruments—has a familiar and formidable presence in the contemporary international landscape of higher education (Chalmers, 2007; Harvey, 2003; Johnson, 2000; Kulik, 2001; Moskal, Stein, & Golding, 2016). Extensive field-based psychometric research has broadly demonstrated the validity and reliability of well-designed student ratings instruments to assess the quality of teaching and the effectiveness of instruction. It has been firmly established that it is possible to effectively limit potential subjectivities in student ratings, such as biases around class size, discipline, grade expectation or instructor characteristics

(Gravestock & Gregor-Greenleaf, 2008; Kulik, 2001; Marsh, 2007). Although it has been suggested that there is limited evidence of specific biases in student ratings, there are incidences of variables necessitating specific responses to be effectively controlled (Benton & Cashin, 2014). Further outcomes have suggested that such variables can be successfully mitigated, either through improved questionnaire design or enhanced methods of survey administration (Marsh, 2007). Yet there are a number of areas—particularly related to the effects of disciplines, teacher performativity and levels of subject difficulty—where some more equivocal outcomes have emerged in research (Aleamoni, 1999; Wachtel, 1998).

Nevertheless, the outcomes of such research has effectively persuaded institutions that student ratings instruments represent a valid and reliable means of assessing the comparative quality of teachers and teaching, programs and assessment, and levels of institutional support. This is demonstrated by the reality that student evaluation data are acting in contemporary institutions as

\* Correspondence address: Facultad de Education, Universidad Alberto Hurtado, Chile.

E-mail address: [stephen.darwin@gmail.com](mailto:stephen.darwin@gmail.com) (S. Darwin).

an increasingly powerful proxy for the quality assurance of teaching, courses and programs across diverse discipline and qualification frameworks (Richardson, 2005). This form of data are also proving increasingly influential in significant institutional judgments about academic appointment, performance and promotion (Darwin, 2016; Harvey, 2003). The global metrics spawned by student ratings systems are also now ubiquitous contributors to the construction of institutional marketing efforts and international university league tables, prospectively influencing the future private (and possibly public) funding of higher education institutions (Herdlein & Zurner, 2015).

Yet the utility of student ratings instruments to positively influence academic teaching remains more equivocal (Beran, Violato, Kline, & Frideres, 2005). For instance, it has been reported that student ratings outcomes remain frequently an unwelcome fringe dweller in the contemporary academy, often responded to with scepticism, uncertainty or unease (Arthur, 2009; Darwin, 2011; Edstrom, 2008). Such scepticism may be amplified as the actual capacity of student ratings to effectively mediate the increasingly complex environments of higher education learning is called into greater question (Schuck, Gordon, & Buchanan, 2008). It has been further asserted that, despite their rising influence within institutions, student ratings outcomes are widely perceived by academics to be inherently narrow and superficial (Edstrom, 2008; Kulik, 2001). Others have argued that student ratings outcomes are perceived to be an inadequate in understanding demanding contemporary expectations on teaching academics to generate high quality learning for ever more heterogeneous, technologically-immersed and demanding student populations (Arthur, 2009; Johnson, 2000; Kember, Leung, & Kwan, 2002).

## 2. Exploring academic perceptions of student evaluation

In order to better understand contemporary teaching academic perceptions of student ratings-based student evaluation, a study was undertaken in a major, internationally recognised university. The study adopted a qualitative methodology, which offered the potential to analysis how the outcomes of student ratings are negotiated within this increasingly complex intersection between individual and collective interests in contemporary higher education institutions (Guba & Lincoln, 1989). Using a critical framework to understand the nature of activity in situated practice, the research sought to respond to two primary research questions:

- a) How do teaching academics understand the role and function of student ratings-based student evaluation?
- b) What is the actual influence of student ratings on academic decision-making around teaching practices and student learning?

Consistent with the qualitative method used, data for the research was triangulated from three distinct sources:

- semi-structured interviews with teaching academics from two teaching programs, all of whom who had received student ratings over the last two preceding semesters ( $n = 32$ )
- two structured focus group discussions involving selected interview participants, that further investigated key themes emerging in interviews
- artefact analysis, involving assessment of key examples of the individual and collective use of student ratings

Respondents participating in the research needed to have substantial engagement with the existing student evaluation system. In addition, the research used a purposive sampling

technique, seeking to capture the experiences of two distinct categories of teachers so us to understand the increasingly dichotomous reality of contemporary higher education environments. The first group of respondents were engaged in teaching within a recently established program, where a relatively small core of tenured academic teachers worked alongside a relatively large peripheral teaching workforce of part-time teachers (many of whom were also engaged in other workplaces). Conversely, the second group of respondents were drawn from a longer established and more core program of the university. As a result, this cohort was made up of a larger tenured teaching workforce and a relatively small group of adjuncts and tutors.

In the semi-structured interviews, teachers and educational leaders of the programs were asked to reflect broadly on their experiences with student ratings. The interviews were framed around three key focus questions:

- What have been your experiences—both positive and negative—with student ratings-based evaluation?
- How useful have you found student ratings as a means of better understanding your teaching effectiveness and the nature of student learning?
- How much influence has the outcomes of student ratings had on your approaches to teaching?

The focus group dialogue was built around the two forums within the respective program-based teams. Here the thematic outcomes of the individual semi-structured interviews were presented for further discussion, clarification and debate. This data was also to become an important catalyst in subsequently formulating what specific action research approaches programs they were to adopt to more effectively harness the expansive potential of the student voice (Darwin, 2011). Finally, key documents surround student ratings were considered, including core tools, policies and reports detailing both individual and collective ratings-based outcomes.

Consistent with the methodology of this study, data collected from respondents and *via* artefacts was interpreted using a broad thematic coding method, which is characterised by Marshall and Rossman (1999) as *emergent intuitive*. This relies on the immersive and intuitive capabilities of the researcher to develop emergent themes for analysis of the data. To effectively manage this, a seven-stage inductive model for thematic analysis was designed. This was framed by the thematic analysis framework for situated research in universities developed by Norton (2009) and integrated the interpretive stages framework advocated by Langemeyer and Nissen (2006).

## 3. Background to the research context

The university in which the research was undertaken was an early adopter of a student ratings-based evaluation model. Student ratings instruments as a voluntary teaching improvement tool for academics were first made available in the early 1980s, coinciding with the introduction of a dedicated academic development unit. As the use of student ratings broadened during the 1990s, they became more accepted as a means as a tool for sparking academic development and improving student retention. During this period, student ratings also gradually became accepted as a legitimate source of evidence to support (or deny) academic appointment, tenure and promotion. At the time immediately preceding the study, the original student ratings system was still in place in its originating form and was enjoying broad use across the institution. However, coincidentally as the research commenced, the university embarked on a major redesign of this long-established system. This reform prosed a series of new strategies around the use of

Download English Version:

<https://daneshyari.com/en/article/4941274>

Download Persian Version:

<https://daneshyari.com/article/4941274>

[Daneshyari.com](https://daneshyari.com)