Appearance, insults, allegations, blame and threats: an analysis of anonymous nonconstructive student evaluation of teaching in Australia

Lakeman, Richard; Coutts, Rosanne A.; Hutchinson, Marie; Lee, Megan; Massey, Deb; Nasrawi, Dima; Fielden, Jann

Appearance, Insults, Allegations, Blame and Threats: An Analysis of Anonymous Non-Constructive Student Evaluation of Teaching in Australia

RICHARD LAKEMAN[ad]

richard.lakemen@scu.edu.au

ORCID: 0000-0002-4304-5431

ROSANNE COUTTS[a]

rosanne.coutts@scu.edu.au

ORCID: 0000-0001-5064-8903

MARIE HUTCHINSON [b]

marie.hutchinson@scu.edu.au

ORCID: 0000-0003-0025-2681

MEGAN LEE[ac]

melee@bond.edu.au

ORCID: 0000-0001-7531-5312

DEBBIE MASSEY[a]

deb.massey@scu.edu.au

ORCID: 0000-0002-0466-1960

DIMA NASRAWI[a]

dima.nasrawi@scu.edu.au

ORCID: 0000-0002-2878-2468

JANN FIELDEN[a]

jann.fielden@scu.edu.au

ORCID: 0000-0003-0084-0838

ᵃ Southern Cross University, Faculty of Health. Southern Cross Drive, Bilinga, QLD, Australia.

ᵇ Southern Cross University, Faculty of Health. Hogbin Drive, Coffs Harbour, NSW, Australia.

ᶜ Bond University, Faculty of Society and Design, University Drive, Gold Coast, Australia

ᵈ Corresponding author

Abstract

Within higher education student evaluations of teaching (SET) are used to inform evaluations of performance of courses and teachers. An anonymous online survey was constructed and implemented using Qualtrics. This study was situated within a more extensive study investigating the impact of narrative SET comments on teaching quality and the health and wellbeing of academic staff. This paper reports specifically on two open questions that were designed to elicit examples of non-constructive and offensive anonymous narrative feedback. Five themes were identified: allegations, insults, comments about appearance, attire, and accent, projections and blame, and threats and punishment. These are represented in non-redacted form. Personally destructive, defamatory, abusive, and hurtful comments were commonly reported. These kinds of comments may have adverse consequences for the well-being of teaching staff, could, contribute to occupational stress and in some cases could be considered libellous. The high prevalence of offensive comments accessible to and shared by teachers may be a reflection of the anonymity afforded to respondents using internet surveys, resulting in deindividuation and enable some respondents to give voice to 'hate speech' which has no place in evaluations of teaching.

Keywords: Student evaluation of teaching; Non-constructive feedback, Anonymous feedback

Appearance, Insults, Allegations, Blame, and Threats: An Analysis of Anonymous Non-Constructive Student Evaluation of Teaching in Australia

Student evaluation of teaching (SET), first introduced in the 1920s, has become a universal feature of higher education (Algozzine et al. 2004). A recent review of studies examining SET (Heffernan 2021) found at least 16,000 higher education institutions routinely survey students internationally. SET has long been valued and utilised in different ways to inform the professional development of teaching staff and course development. The authors of early reports on SET, suggest that evaluations needed to be anonymous so that they "could not be used against students", that results should only be available to teaching staff to use for their purposes, and noted that a large proportion of comments, whilst mostly favourable, were typically about the personality or attire of teaching staff (Ostlund 1955). However, since the adoption of "new managerialism' in the higher education sector (from the mid-1990s), SET has been institutionalised, procedures standardised, anonymous numerical rating of teacher competence as well as satisfaction with courses normalised, and teaching staff "benchmarked' against others (Lee, Coutts, et al. 2021, 7). As a 'key performance indicator' and the primary means of evaluating teaching effectiveness, SET is now routinely used to inform judgments about promotion and tenure (Clayson and Haley 2011). A key element of contemporary SET in Australia is that responses are completed anonymously and online. It is proposed that anonymity is a factor that leads to antisocial behaviour. This paper focuses on anonymous narrative feedback and, comments that are perceived as non-constructive or offensive.

Background

SET typically involves evaluation of teaching staff and courses, with students being invited to rate both, usually using Likert scales and via anonymous narrative comments.

4

While some have argued that a multidimensional approach to evaluating teaching effectiveness, including these factors, is valid and reliable (Burdsal and Harrison 2008), these assumptions are increasingly being challenged (Heffernan and Bosetti 2021; Hornstein 2017; Esarey and Valdes 2020). For example, a recent meta-analysis concluded that SET ratings are unrelated to student learning and that students do not learn more from teachers with higher SET ratings (Uttl, White, and Gonzalez 2017). Nevertheless, teacher and course ratings accompanied by anonymous feedback are the norm and facilitated by near-universal connectivity to the internet and the ease with which questionnaires can be deployed on mobile devices at any time. Little is known about the motivation of students who complete SET relative to those who decline. Suffice to say, overall response rates are typically low; students complete evaluations at a time of heightened anxiety (usually during times when final assessments are due or have been marked) and often in anticipation of doing poorly (Heffernan 2021).

The institutionalisation and massification of anonymous SET has accompanied the availability and uptake of technology. SET is routinely undertaken online and anonymously at a time convenient to students. Since the famous Stanford prison experiments, anonymity has been considered one of the antecedents that can lead to a lowered threshold for expressing otherwise inhibited behaviours (Diener 1977). Anonymity amongst other variables leads to a state of deindividuation characterised by a loss of self-observation, self-evaluation and concern for social evaluation (Diener 1977). Deindividuation or the faceless anonymity of the internet is theorised to give rise to online exploitation, cyber-bullying, hate speech, and general meanness in some online interactions (Myers 2016). The anonymity that the internet affords is associated with undisciplined, disinhibited and anti-normative behaviour (Guo and Yu 2020), that would not be countenanced in face-to-face interactions. In light of this it is likely that anonymous internet facilitated SET increases the likelihood of non-constructive

5

feedback, bullying and incivility which appears to be increasing in higher education (Heffernan 2021; Heffernan and Bosetti 2021).

These factors and others such as the psychological state and motivation of the student completing the SET (Hoel and Dahl 2019), the time of completion, and not being in the physical presence of those being commented on or other cues for social etiquette may lead to ratings and comments which are not accurate and sometimes counter to social norms. Even the now seemingly quaint and old fashioned idea of requesting SET be completed anonymously 'in class' has been found to double the completion rate although the impact on the tone of narrative feedback is unclear (Kuch and Roberts 2019).

It is mainly in providing for SET in or at the end of class that anonymous versus open SET has been examined. Afonso et al. (2005) noted that in a University academic medical centre context SET scores were significantly lower on all items when a questionnaire was administered anonymously. They concluded that anonymous evaluation was a more accurate reflection of teaching performance. This may be true in high esteem courses, for professions that have traditionally rigid hierarchies, and where challenging the consultant or expert, who may have a relationship with the student over many years, is difficult. It is far less clear whether anonymous SET serves any useful function in the now massified and competitive higher education marketplace.

Small classes and long-term relationships with teaching staff throughout a degree were the norm when anonymous SET was conceived (and less than 3% of the population completed degrees). Since the beginning of the internet age, the majority of school leavers (>70%) now will complete degrees in most OECD countries and the projections are that this percentage will increase (Teichler and Bürger 2008). The timing of SET, typically before grades are released but after teaching staff can make any adjustments in response to feedback that might benefit respondents, nullifies the argument that students are potentially vulnerable

6

when undertaking SET and therefore need to do so anonymously. Recent studies have found that SET after grade release is influenced by the grade point average (GPA). Increasing the GPA will lead to a significant increase in SET ratings (Berezvai, Lukáts, and Molontay 2021; Stroebe 2020).

One review of the literature found that students at times falsify SET and a majority of students in a survey knew of respondents who provided false or misleading responses to questions (Clayson and Haley 2011). Other researchers suggest that the perceived charisma of the lecturer accounts for most of the variation in ratings of lecturer ability and most course attributes (Shevlin et al. 2000). While ratings have been examined in detail, the anonymous narrative feedback has not been scrutinised systematically and across multiple sites. Tucker (2014) examined over 30,000 comments from 17,855 surveys at one Australian University in 2010 and found only thirteen abusive comments and 46 unprofessional comments. This number was perceived to have remained static and was perceived as so insignificant that removing anonymity was not considered. Tucker (2014) reported that comments were only shared with the course coordinator and head of school at that University. Others have suggested that in recent years the rates of unconstructive, and offensive comments have increased (Heffernan 2021; Heffernan and Bosetti 2021). The actual frequency and prevalence of non-constructive and/or offensive narrative comments is not reported in the literature to date. It is also rare for actual examples of non-constructive or offensive comments to be shared in the literature and this paper addresses this gap.

Aims

This project aimed to collate and categorise examples of non-constructive SET and is located within a larger project that aims to survey the impacts of anonymous narrative feedback on teaching quality and health and well-being of academic teaching staff. It was

7

undertaken in the context in which the workforce is increasingly casualised and insecurely tenured and when ongoing tenure is often highly dependent on positive reviews of teaching (Lee, Coutts, et al. 2021). Many commencing academic teaching staff will have had little prior experience of receiving anonymous feedback. This paper also aims to represent in an anonymous but non-redacted form the kind of non-constructive feedback teachers may expect to receive at some time in their career. Thus, early-career higher education teaching staff may be better psychologically prepared for non-constructive feedback and consider collating a broader range of evidence (e.g., peer observation) to support their performance reviews.

Methods

This research utilised a mixed-methods approach to retrospectively investigate the responses of teaching academics to anonymous narrative feedback from students across Australian universities. An anonymous online survey was constructed and deployed using Qualtrics. The questionnaire was comprised of thirty questions focusing on employment type, history, and demographics (13 questions), processes employed to share anonymous feedback and examples of constructive and non-constructive feedback (7 questions), perceptions of the impact on teaching quality (5 questions) and impacts on social and psychological wellbeing (5 questions). This paper reports on the two open questions eliciting examples of non-constructive and offensive anonymous narrative feedback. The research team extensively trialled the questionnaire before deployment and approval obtained from the Southern Cross University Human Research Ethics Committee (2021/047).

Higher education teaching staff in Australia were recruited via a snowballing strategy and social media over three months in 2021. This included emails to colleagues, promotion on twitter and facebook as well as an invitation to participate in 'The Conversation' (Lee, Nasrawi, et al. 2021). Data were analysed using SPSS v28. Respondent characteristics were

8

analysed through univariate analysis (frequency, percent, means and standard deviation). The Chi-Square test of independence was employed to establish significant relationships between categorical variables. The responses to open-ended questions were thematically analysed by the first author with the assistance of the software package NVIVO, and following the principles of inductive content analysis outlined by Corbin and Strauss (2014). This was an iterative process in which every line and example was coded, and as these appeared to coalesce into sub-themes, these were labelled and described, previous data coding reviewed, and examples subsumed under the new themes. The coding structure was reviewed by the research team with decisions made about data reduction and agreement on final themes. All variation in responses was captured in the narrative findings under the emerging themes. The verbatim findings are presented as quoted, including profanity and with grammatical and spelling errors intact.

Findings

The Sample

The total number of respondents was 791. The majority were employed outside the Group of 8 universities (81%, n=641), that is those considered Australia's eight leading research-intensive universities. Most reported working in regional centres (57.4%, n=454) rather than metropolitan settings (42.6%, n=337). Most were female (77.6%, n= 614), followed by male (20.9%, n= 165) or non-binary (0.4%, n=3). The mean age of the cohort was 48.6 years (SD 8.24). The mean duration of employment in the university sector was 13.44 years (SD10.1). Ethnicity, employment status and grade of respondents is reported in table one.

Table 1: Additional demographic details

Almost all reported receiving non-constructive comments about their teaching and courses at some time (91.4%, n = 723). No dependent relationship was identified between the receipt of distressing, offensive or disrespectful student feedback and, gender (χ^2 (16, $N = 675$) = 21.17, $p = .172$), ethnicity (χ^2 (7, $N = 697$) = 4.840, $p = .679$) or employment grade (χ^2 (4, $N = 645 = 3.081$, $p = .544$). Close to one-third of respondents reported that non-constructive, personalised, or offensive comments were not redacted before being shared with the staff member concerned or others (27.3%, n = 216). As illustrated in table two, only a minority reported that anonymous narrative feedback was shared exclusively and as a matter of policy with the staff member concerned (although, it is unclear who has access to these data). Close to half reported this feedback was shared by immediate supervisors and others in their workgroups. Thus, for most respondents, anonymous feedback needs to be acknowledged as contributing to a semi-public discourse about them as individuals. As was often made clear by some respondents, this feedback was influential in shaping their career trajectories and promotional prospects.

Table 2: At your Current University how Widely is Anonymous Narrative Student Feedback Shared?

Respondents were asked to provide examples of non-constructive feedback and 552 (68%) provided examples (16,484 words), which mostly took the form of direct quotes. However, some respondents also provided a commentary. Respondents were also asked if they had ever received anonymous narrative feedback about their teaching or courses which they found distressing, offensive or disrespectful and 68.8% (n=544) responded yes. Of these respondents, 378 provided further examples of offensive or distressing comments (9,867 words). There was some repetition in the comments about distressing, offensive or personalised feedback so only new statements were included.

The Threads and Themes

Examples of feedback considered by respondents to be nonconstructive fell into five themes: (i) allegations (ii) insults;(iii) comments about appearance, attire, and accent (iv) projections and blame, and (v) threats and punishment (see Figure 1). Although conceptually, each of these themes represented 20% or more of the examples provided, there was considerable overlap. For example, comments about appearance, dress, or accent were often couched as insults, and many examples provided encompassed multiple themes. Four threads were interwoven in the comments which reflected poles on continua but were also independent sub-themes. Irrelevant and 'feedback' unrelated to evaluation was at one pole benign and quite transparent and on the other pole hateful and defamatory. There was also the superficial and absurd which could be dismissed as such but on the other end of the continuum were catastrophic and totalising comments, which are impossible to address because they reflect a dichotomous view of both the recipient of feedback, the course, or University.

This intersection of themes and subthemes was illustrated in numerous quotes of diatribes or paragraphs of invective language e.g.

> *X should be sacked she is not up to the job they should replace X with the*
>
> *lab techs they know more than her… X is way to old to be doing this job*
>
> *this course is a joke I'm going to another Uni… this unit is the worst I have*
>
> *ever done… I will tell all my friends to take valium before they attend it*
>
> *next year…*

Some comments verged on the incoherent, although the vitriolic sentiment was apparent. For example:

11

This subject is a disgrace to the University and to the students who have to listen to it. I suggest you practice what you preach. For example when you have priests that rape children then this is a slap in the face to any victim forced to do this pathetic subject.

Figure 1: A Typography of Non-Constructive Feedback

Allegations

Many examples were personalised allegations of impropriety, substandard behaviour or allegations of flawed character (see table three). These allegations were often offered with no evidence to support statements:

It was clear lecturer x had no idea what she was talking about and didn't want to be there…

Closely related were allegations of incompetence such as not knowing how to run a course. These statements also tended to overlap with other themes of projection, and blame as the following example demonstrates (the capitalisation is preserved):

X teaches stuff that makes no sense and is no use, none of it makes any sense and she makes us feel like idiots (sic)

The next most common cluster of allegations was around being *rude, racist, ignoring students*, or being disrespectful towards students. These were often framed as insults. "*She is really rude which is why everyone hates her*". Perhaps more serious were allegations of academic misconduct such as "*Deliberately sabotaging students' ability to learn by hiding content*" although like most allegations, these were unverifiable and, as many respondents noted, were baseless and false:

12

My tutor is clearly racist. The reason for my failure in this course lies

entirely at my tutor's feet. My tutor should be reported for academic

misconduct. She discriminates and has favourites, that is why I failed.

Many examples were given of allegations that people were harsh, unfair, nepotistic or discriminatory whilst also alleging that they were unavailable, unfair or unhelpful.

Table 3: Personalised Allegations or Accusations

Appearance, Attire and Accent

A large number of examples and comments related to appearance (n = 66), attire (n = 31) and voice or accent (n = 45). These were often framed as insults and were gender-specific with women most frequently reporting commentary on multiple aspects of their appearance and presentation. Many simply reporting that comments on appearance, clothing, hair and dress were common:

X is a disgrace to the university and X department. She is rude to students when they

ask for help in tutorials and does not care for our learning. She presents herself in an

unprofessional and frankly disgusting way, her clothes are extremely unflattering and

dresses way too short. I have been 'flashed' by her on 3 occasions, where she sits with

her legs wide open in minuscule dresses covers nothing. I can't even look at her

because she makes me feel sick.

Most of the examples related to appearance and attire were clearly intended to be insulting and some respondents reported them as being "*nasty*" and hurtful although some suggested they were intended as jokes and others labelled comments as examples of sexual harassment *"nice booty", "I love your tits",* or "*she shouldn't wear such sexy clothes, it's*

13

distracting". One person stated that a student used anonymous feedback to make a proposal of marriage.

Comments on appearance included that the lecturer "*had dandruff*", "*needed to pluck their eyebrows*", "*looked good*", "*looked like an animal*", "*looked scary*", looked like "*something the cat dragged in*", and had "*weird body movements*". Many comments were highly insulting: "*Fat pig*", "*Fat bitch*" or "*Too ugly to teach*"; or they were part of a diatribe of criticism:

> *X is useless, always has a cup of coffee with her and looks like she never*
>
> *sleeps. If she stopped drinking coffee maybe she would be more intelligent.*

Comments on attire were gendered and statements about students not liking people's clothing: "*She dresses like a hag*". One respondent stated that comments about the way she dressed were circulated to the entire school. Others noted that students complimented them on their jewellery, shoe collection, choice of clothing or other irrelevant details.

A large number of comments addressed people's voice and most frequently that students found it "*boring*", *"annoying", "irritating"* or *"droning",* that people spoke too fast or slow, or couldn't annunciate correctly: *her voice is annoying and makes me sick/gives me a headache*. Some might have been constructive in as much as people's paralinguistic cues and idiosyncrasies which are under their control might be modified, but often these were framed as witticisms or insults: *Says yeah a lot. This term she has said yeah 2,362 times.* Comments on accent were also common and had racist overtones, with many students stating that they could not understand teaching staff because of their accent despite them being native English speakers: *She needs to adopt an Australian accent better.*

14

Insults & Invective

An insult regardless of veracity is disrespectful, scornful or abusive. Most comments cited as non-constructive, or offensive appeared intended to wound or disrespect the person they were aimed at. They were also by and large unverifiable, defamatory and often quite absurd. Most insults were not particularly imaginative and involved name-calling such as: *"bitch", "bitter", "crap", "cunt", "devils spawn", "dick", "dog", "dinosaur", "idiot", "loser", "lacklustre", "mentally unstable", "missing in action", "mole", "Nazi", "needs to chill", "out of control", "pathetic", "psychotic", "senile", shit" "smiling assassin", "stick with the day job* [medicine]", *"time-waster", "TRASH", "unhappy", "useless", and "gangsta tutor"* (the latter also perplexed the respondent). There was also a considerable variation on the theme*: "I hate everything about you", "the teacher is the worst ever,"* or *"everyone else is better".*

One respondent stated that the expletives were blocked out at their University, but the meaning was clear. Another stated that at their University, expletives and offensive comments are redacted but rephrased which left them wondering how bad the original 'feedback' must have been:

> *[student found lecturer repetitive and did not find the delivery interesting],*
> *or [student expressed their belief that the lecturer had significant bias]. It*
> *feels just as bad. How nasty must they have been for them to have to redact*
> *it like that?*

Many respondents stated that they could not repeat some of the things that had been written about them, or they simply no longer look at their SET results for fear of what they might say. Many people did report that they found the insults wounding and hurtful:

15

> *A student wrote a comment that was very hurtful about me eating a bat and*
> *giving COVID-19 to the world. That I was personally responsible for the*
> *death of millions of people.*

Others were equally imaginative, including that the lecturer: "*needed moderation*", that the lecturer: "*had killed their love of the subject*", "*the doctor is lacking in confidence and should sign up to a Toast Masters Course*" or that attendance at a lecture was "*cruel and unusual torture*". More often than not potentially cutting witticism degenerated into puerile name calling:

> *You are a cultural Marxist, your Wokeness undermines everything you do.*
> *Not all your students are left wing nut jobs like you. You seriously need to*
> *lose some weight.*

Others were highly personal and reflected personal knowledge about the teacher's circumstances:

> *What the fuck did you think you were doing to take a couple of days off for*
> *your grandmother's funeral when we had an assignment due?*

Under the veil of anonymous feedback students appeared to have license to express overtly racist, ageist, sexist and homophobic insults with impunity: "*I'd do her*". Some racist comments were expressed as allegations, or exhortations for teachers to be more Australian but others were hateful:

> *She is a fucking wog who needs to stop shopping at Kmart ...you fat bitch*

Insulting comments were made about people's age (either too young or old to teach) and these were frequently gender specific:

The lecturer is like a lady who lives alone and has a cat.

Stupid old woman needs a good fucking

Respondents who identified as non-binary or LGBTQI+ noted that their sexuality was a source of derision at times. Homophobic insults were also cited by others as examples of non-constructive anonymous feedback:

Why don't you just come out of the closet? I pity your wife and children.

Projection and Blame

Arguably all the examples of non-constructive feedback were irrelevant and unrelated to the quality of teaching and learning. Some respondents spoke of students utilising feedback to express grievances about what they did not like but beyond the teachers' control. For example, that students stated they had '*wasted their money*", did not like the class size, or thought courses should be run by clinicians rather than academics. One respondent explained how students would sometimes externalise their behaviour:

Another common theme is externalising e.g. "I was made to buy a text-book

and I didn't use it" - like it's the staff member's fault they didn't open it?

Most respondents reported being aware that some disgruntled or aggrieved students projected their grievances and that students would never make similar statements to their face. Several noted that students sometimes colluded in orchestrating a barrage of negative comments and reported the perception that some students tried to "*destroy"* them. The sub-theme of 'catastrophising and totalising feedback' was particularly prevalent in this theme with large numbers of students being reported as expressing in feedback that the content was "*irrelevant", "unnecessary", or "uninteresting*"; the course "*too hard*", "*long*" or "*complex*"; or assessments simply "*wrong*". Reports of feedback included but were not limited to: "*I*

17

learned nothing", "I didn't know what to do", "I taught myself" or more commonly that the ineptitude or the vindictiveness of the tutor led to their failure:

That fucking dyke bitch failed me she's fucking useless that's why I failed.

The most common examples of projection and blame raised by respondents were also examples of false or misleading statements, which were likely tendered as examples because of the injustice of these false statements remaining on the semi-public record (often in direct contradiction of other feedback). Such projected and false statements were often about the person "*never replying"* or "*ignoring emails*", "*not returning to work on time",* and not being explicit about course requirements. It was clear that many students were not happy with their grades or remark and tended to project blame onto teaching staff:

The teachers in this course are racist because all the Asian students get low

grades

Threats of Retribution and Calls for Punishment

It is clear that some students blame teaching staff for their, failures, or problems unrelated to that have little to do with the performance of the teacher being evaluated. Many respondents (n>50) reported being told they should be: *sacked*. These were typically part of a totalising and catastrophizing discourse whereby the student reported that everything was awful and in response to the question "How can this topic, course be improved?" the response was to: *sack the teacher*. Some respondents elaborated on the dynamics that led to such 'feedback':

A student (who was easily identifiable) who said I should be sacked

because I wouldn't assist him by printing hundreds of pages of material for

him. The same student, misunderstanding Socratic method teaching,

> *claimed that I deliberately asked him questions that I knew he didn't know*
>
> *the answers to (apparently presuming that I could read his mind).*

Variations on demands for people to be sacked were exhortations that the teacher should resign, not practice in their field, that they need retraining, that they are awful or unsuitable, that the University should be embarrassed, that the teacher should not teach that subject, or that they should be replaced. These were sometimes made with threats that the student would withdraw if the teacher continued. One reported that students had threatened legal action or going to the media because they did not like the grade they received.:

> *You're fake you act sincere, but in reality you're a phony, a hypocrite,*
>
> *double-dealing and pretentious. To be honest you are the worst tutor that I*
>
> *have come across. No likes to deal with a two face. Stop teaching your*
>
> *embarrassing yourself. Quit while you're ahead. I hope you get fired soon.*
>
> *Why did you take the job your clearly not capable? Why don't you just*
>
> *resign?*

Some threats were considerably more menacing. One reported that a student had written: *I'd like to shove a broom up [the lecturer's] arse* and another that: *she should be stabbed with a pitchfork*. Some of these threats were puerile: "*X should shut up and die*" or "*If I was X, I would jump off the tallest building and kill myself if I was that dumb*". However, threats such as this one to *watch out* were couched in vitriolic diatribes:

> *This bitch should be fired immediately. Why is someone this ugly allowed*
>
> *to teach? She better be careful I never see her in the car park. She needs to*
>
> *get a better fashion pick. Her clothes are hideous.*

19

Discussion

This research aimed to establish the occurrence and type of non-constructive anonymous narrative comments in SET in Australian Universities. Staff across Australian universities shared a vast array of feedback from students. This was able to be categorised as either-or combinations of insult, defamatory statements, and a plethora of words or phrases which apportioned blame, criticism, accusation, and threat. That, many respondents directly quoted or copied and pasted from SET they had received was a surprise to the research team. However, this might also be in part due to how readily at hand these data are in this relatively new era of online teaching and administration.

The primary limitations of this research relate to the snowballing method of recruitment and the diversity of experiences in Australian teaching staff may not have been adequately captured. The researchers were health professionals and by and large engaged with other health professionals although at least one third of respondents were not health professionals. The majority of respondents were women which could explain the clearly gendered examples of non-constructive feedback. However the findings echo the Heffernan (2021) review that argued student surveys were flawed, prejudiced against those being assessed and also included increasingly abusive comments directed mostly towards women and marginalised groups. There is also likely to be self-selection bias in the non-probability sampling (Bethlehem 2010) meaning that although the majority of people in this sample reported receiving non-constructive feedback and also found it offensive or distressing, these findings cannot be extrapolated to all academics in Australia.

Few recent studies have explicitly examined the content of anonymised SET. Mowatt (2019) has provided one of very few published reflections on the experience of receiving SET (over 12 years teaching) and described it as a form of "racialised intellectual violence".

20

Mowatt (2019) noted that men of colour consistently receive the lowest ratings in SET and provided examples of non-constructive feedback. Dunegan and Hrivnak (2003, 283) have noted that it is rare for teacher evaluations to be shared with students who receive virtually no feedback about the outcomes of SET which potentially leads to "mindless teaching evaluations". However, mindless and invalid SET (which might be addressed by sharing feedback liberally with students) is unlikely to address the problem of some students purposefully intending to hurt and defame. Mowatt (2019) clearly articulated the purpose of the comments routinely received: "They serve a purpose—for my removal". Mowatt (2019) further asserts (and we are in agreement) that it is naïve to consider that students are unaware of the importance that administrators hold anonymised feedback. The respondents of this survey seem to be well aware that whilst it might not advantage individual students directly to malign or defame academic teaching staff there appeared to be quite clear motives to punish, malign or hurt academics. Anonymous feedback can impact on teaching staff emotionally and professionally. This will be explored in other papers arising from this project.

Further research should explore the current prevalence of non-constructive feedback and identify if the prevalence of non-constructive comments has increased over time. In our sample, examples of non-constructive feedback were readily recalled by respondents as were clearly defamatory statements, which suggests possible negative impacts on their wellbeing. Some respondents stated that they had not received negative comments for years, and this may reflect professional development career progression and reduction in teaching responsibilities or, accommodations made in response to, or in anticipation of potential wounding comments. At best the examples we have provided suggest any associated numerical rating of teachers associated with non-constructive comments ought to be considered invalid. However, it appears that even if written comments are redacted in some

21

way that numerical ratings are preserved.

The findings from this survey appear at odds with Tucker's (2014) review of feedback at one Australian university in 2010 which found few examples of offensive feedback. Tucker (2014) did not provide verbatim examples of non-constructive feedback as we have done. However, it can be readily inferred that this sizeable sample of academics in this sample had at hand many examples of offensive non-constructive comments. It is likely that the findings from this survey reflect an acceleration of a trend towards deindividuation (Diener 1977) and antisocial behaviour in some students.

At the time this survey was deployed the majority of Australian universities were teaching online due to the COVID-19 pandemic which had wide reaching impacts on students and staff. Universities shed a large number of largely casual workers and the public discourse supported by Government policy was that the value of universities was primarily about delivering 'job-ready graduates' reinforcing a view of students as customers (Lee, Nasrawi, et al. 2021). At the same time, the experience and impact of 'cyber-bullying' (which some of the examples of non-constructive feedback clearly reflect) has increased in public awareness (Jenaro, Flores, and Frías 2018) and arguably has become a common experience. There are now concerted efforts to understand and reign in 'cyber-bullying' in workplaces (Herron 2021) and to hold social media conglomerates to account (Ullmann 2021). However, it appears that universities (at least in Australia) are one of the last remaining bastions that allow and enable students to say anything, regardless of veracity or impact about their teachers with impunity.

We were shocked by the examples of feedback provided. The quantity of vitriol, personalised, offensive and hateful comments was not expected and led to some collective reflection about the personal impacts of reading anonymous feedback of this nature even when not about oneself personally. The majority of respondents (88%) noted that SET

comments were circulated to others beyond their immediate supervisor. While many non-constructive examples were superficial, absurd, irrelevant and unrelated, many were also defamatory and hateful. If such comments took the form of graffiti, they would be quickly removed in the interest of good taste and, if published in a public sphere about named individuals, might be considered libellous.

Jones, Gaffney-Rhys, and Jones (2014) discussing the risks of publishing SET data in the United Kingdom noted that students are afforded protection by making anonymous comments and expressing opinions but the publication of statements of fact which are untrue and lead to adverse consequences for individuals is potentially libellous. It is the anonymity and protection this SET process affords which allows, enables, and possibly encourages these non-constructive comments about an academics teaching. It is hard to justify continuing to enable students to undertake SET anonymously and it is unclear what of value would be lost if students were to provide potentially identifiable feedback.

Conclusion

Our findings evidence the nature of insulting and offensive commentary collected through electronic student evaluation of teaching in Australian universities. Although the research has been conducted in one country, the findings are consistent with concerns about the validity and usefulness of student evaluation that have been widely canvassed. What has been given less considered attention is the use of online student evaluation of teaching as a platform to harass, offend, and at times, menace teachers in higher education. Whether the threshold has been reached for this behaviour to be considered abuse remains to be established. This is particularly relevant when the commentary is personalised and clearly intends to cause harm or hurt. We contend action is required to assure the accountability and workplace safety of these systems and to curb increasing incivility in higher education.

23

References

Afonso, N.M., L.J. Cardozo, O. Mascarenhas, A. Aranha, and C. Shah. 2005. "Are anonymous evaluations a better assessment of faculty teaching performance? A comparative analysis of open and anonymous evaluation processes." *Fam Med* 37 (1):43-7.

Algozzine, B., J. Gretes, C. Flowers, L. Howley, J. Beattie, F. Spooner, G. Mohanty, and M. Bray. 2004. "Student Evaluation Of College Teaching: A Practice In Search Of Principles." *College Teaching* 52 (4):134-41. doi: 10.3200/CTCH.52.4.134-141.

Berezvai, Z., G.D. Lukáts, and R. Molontay. 2021. "Can professors buy better evaluation with lenient grading? The effect of grade inflation on student evaluation of teaching." *Assessment & Evaluation in Higher Education* 46 (5):793-808. doi: 10.1080/02602938.2020.1821866.

Bethlehem, J. 2010. "Selection Bias in Web Surveys." *International Statistical Review* 78 (2):161-88. doi: https://doi.org/10.1111/j.1751-5823.2010.00112.x.

Burdsal, C.A., and P.D. Harrison. 2008. "Further evidence supporting the validity of both a multidimensional profile and an overall evaluation of teaching effectiveness." *Assessment & Evaluation in Higher Education* 33 (5):567-76. doi: 10.1080/02602930701699049.

Clayson, D.E., and D.A. Haley. 2011. "Are Students Telling Us the Truth? A Critical Look at the Student Evaluation of Teaching." *Marketing Education Review* 21 (2):101-12. doi: 10.2753/MER1052-8008210201.

Corbin, J., and J. Strauss. 2014. *Basics of Qualitative Research: Techniques and procedures for developing grounded theory*. 4th ed: Sage.

Diener, E. 1977. "Deindividuarion: Causes and consequences." *Social behavior and personality* 5 (1):143-55. doi: 10.2224/sbp.1977.5.1.143.

Dunegan, K.J., and M.W. Hrivnak. 2003. "Characteristics of Mindless Teaching Evaluations and the Moderating Effects of Image Compatibility." *Journal of Management Education* 27 (3):280-303. doi: 10.1177/1052562903027003002.

Esarey, J., and N. Valdes. 2020. "Unbiased, reliable, and valid student evaluations can still be unfair." *Assessment & Evaluation in Higher Education* 45 (8):1106-20. doi: 10.1080/02602938.2020.1724875.

Guo, K.H., and X. Yu. 2020. "The anonymous online self: Toward an understanding of the tension between discipline and online anonymity." *Information Systems Journal* 30 (1):48-69. doi: https://doi.org/10.1111/isj.12242.

Heffernan, T. 2021. "Sexism, racism, prejudice, and bias: a literature review and synthesis of research surrounding student evaluations of courses and teaching." *Assessment & Evaluation in Higher Education*:1-11. doi: 10.1080/02602938.2021.1888075.

Heffernan, T., and L. Bosetti. 2021. "Incivility: the new type of bullying in higher education." *Cambridge Journal of Education*:1-12. doi: 10.1080/0305764X.2021.1897524.

Herron, M.M. 2021. "Social Media Bullying in the Workplace: Impacts on Motivation, Productivity, and Workplace Culture." In *Handbook of Research on Cyberbullying and Online Harassment in the Workplace*, 72-89. IGI Global.

Hoel, A., and T.I. Dahl. 2019. "Why bother? Student motivation to participate in student evaluations of teaching." *Assessment & Evaluation in Higher Education* 44 (3):361-78. doi: 10.1080/02602938.2018.1511969.

Hornstein, H.A. 2017. "Student evaluations of teaching are an inadequate assessment tool for evaluating faculty performance." *Cogent Education* 4 (1):1304016. doi: 10.1080/2331186X.2017.1304016.

Jenaro, C., N. Flores, and C.P. Frías. 2018. "Systematic review of empirical studies on cyberbullying in adults: What we know and what we should investigate." *Aggression and Violent Behavior* 38:113-22. doi: https://doi.org/10.1016/j.avb.2017.12.003.

Jones, J., R. Gaffney-Rhys, and E. Jones. 2014. "Handle with care! An exploration of the potential risks associated with the publication and summative usage of student evaluation of teaching (SET) results." *Journal of Further and Higher Education* 38 (1):37-56. doi: 10.1080/0309877X.2012.699514.

Kuch, F., and R.M. Roberts. 2019. "Electronic in-class course evaluations and future directions." *Assessment & Evaluation in Higher Education* 44 (5):726-31. doi: 10.1080/02602938.2018.1532491.

Lee, M., R. Coutts, J. Fielden, M. Hutchinson, R. Lakeman, B. Mathisen, D. Nasrawi, and N. Phillips. 2021. "Occupational stress in University academics in Australia and New Zealand." *Journal of Higher Education Policy and Management*:1-15. doi: 10.1080/1360080X.2021.1934246.

Lee, M., D. Nasrawi, M. Hutchinson, and R. Lakeman. 2021. "Our uni teachers were already among the world's most stressed. COVID and student feedback have just made things worse." *The Conversation*.

Mowatt, R.A. 2019. "Twelve Years a Servant: Race and the Student Evaluation of Teaching." *SCHOLE: A Journal of Leisure Studies and Recreation Education* 34 (2):109-19. doi: 10.1080/1937156X.2019.1622949.

Myers, D.G. 2016. "A social psychology of the internet." *International Forum of Teaching and Studies* 12 (1):3-9.

Ostlund, L.A. 1955. Anonymous Student Evaluation of College Teaching. Paper presented at the Proceedings of the Oklahoma Academy of Science.

Shevlin, M., P. Banyard, M. Davies, and M. Griffiths. 2000. "The Validity of Student Evaluation of Teaching in Higher Education: Love me, love my lectures?" *Assessment & Evaluation in Higher Education* 25 (4):397-405. doi: 10.1080/713611436.

Stroebe, W. 2020. "Student Evaluations of Teaching Encourages Poor Teaching and Contributes to Grade Inflation: A Theoretical and Empirical Analysis." *Basic and Applied Social Psychology* 42 (4):276-94. doi: 10.1080/01973533.2020.1756817.

Teichler, U., and S. Bürger. 2008. "Student enrolments and graduation trends in the OECD area: What can we learn from international statistics." In *Higher Education to 2030: Demography*, 151-72. Organisation for Economic Co-operation and Development: Centre for Educational Research & Innovation.

Tucker, B. 2014. "Student evaluation surveys: anonymous comments that offend or are unprofessional." *Higher Education* 68 (3):347-58. doi: 10.1007/s10734-014-9716-2.

Ullmann, O. 2021. "Time to Rein in Social Media?" *The International Economy* 35 (2):8-11.

Uttl, B., C.A. White, and D.W. Gonzalez. 2017. "Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related." *Studies in Educational Evaluation* 54:22-42. doi: https://doi.org/10.1016/j.stueduc.2016.08.007.