You're better than you think you are:

A Comparison of lecturer and student rated evaluations of teaching on a Professional Doctorate training programme

Dr Bob Patton

Lecturer in Clinical Psychology

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27619v1 | CC BY 4.0 Open Access | rec: 28 Mar 2019, publ: 28 Mar 2019

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.27619v1 | CC BY 4.0 Open Access | rec: 28 Mar 2019, publ: 28 Mar 2019

Contents

Overview.

Student evaluation of teaching content and delivery has been undertaken at the University of Surrey for many years. Typically students complete Module Evaluation Questions (MEQs) at the end of each teaching module. Feedback from the MEQs is provided to teaching staff and this contributes towards individuals' assessment of their appraisal targets, and to course convenors to help the ongoing development of course materials and staff development. Trainees attending the PsychD Clinical Psychology Doctorate programme complete a "Feedback on Learning" questionnaire (based on the standard university MEQ) on a weekly basis for each lecture they attended. While the evidence base for student assessment of leaning is well founded, the university does not at present seek to understand lecturers own reflection on their delivery of teaching. By determining differences and commonalities between educators and students experiences of teaching, it is hoped that we may identify areas for further staff development and training, and to facilitate regular reflection of teaching performance from staff. This type of research is in line with the UK Professional Standards Framework, with reference to its' emphasis on ensuring good practice and excellent student experience, demonstrating an evidence based approach to scholarship and learning, utilisation of methods to evaluate the effectiveness of teaching, and the provision of assessment / feedback to learners (where we view in this case the teachers as learners themselves as part of a constructionist paradigm).

Background.

Pedagogical research has demonstrated that student feedback questionnaires are a valid and reliable method of gaining feedback on learning and teaching, with student rating of performance correlating strongly with independent assessment of teaching quality (Marsh, 1987). Receiving such feedback can have a positive effect in improving teacher performance. In a meta-analysis of 22 studies exploring effectiveness of feedback, researchers observed a significant improvement in instructor ratings for those who received feedback (Cohen, 1980). The process of critically evaluating teaching can both help learners to consider how learning is being delivered as well as allowing teachers to develop an insight into how their performance is perceived by their intended audience (Seldin, 1997). In order to be effective, feedback should be relevant, immediate, factual, helpful, confidential and respectful (Ovando, 1994). Student perceptions of the quality of their teaching also contribute towards their overall educational experience, and in turn this can mediate their own approach to study and influence their own learning outcomes, with students who have a positive experience of teaching utilising a 'Deep' approach to learning and achieving superior grades to those who perceived teaching as low quality (Lizzo, Wilson, & Simons, 2002).

Does the use of student feedback improve overall teaching quality? The provision of feedback alone is unlikely to have any significant effect upon teacher performance unless teachers themselves take action to address any criticism or weakness identified. Given the additional workload imposed on staff and students alike in engaging with the feedback process, evaluation of its effectiveness is a requirement. Teachers own attitude towards the assessment of their performance is likely to influence their response to feedback, in particular if they perceive their performance to be adequate, they are unlikely to respond to criticism (Seldin, 1995). In a long term study of the impact of student feedback on teaching quality, Hong Kong University found that only 4/25 departments demonstrated significant changes in six

specific indices of teaching quality over a four year period, and that three of these changes were negative (Kember, Leung, & Kwan, 2002). In this case feedback had been implemented without consultation with the wider student body and staff, and this may well have detracted from its effectiveness. This negative impact of 'imposed' assessment of teaching was also observed in a study that found that over a four semester period, where feedback was provided at the end of each semester, an initial improvement in quality from semester one to two was followed by a decrease from the second to fourth semester (Lang & Kersting, 2007). This would suggest that it is important for the feedback process to be presented in context and that engagement from all parties is necessary.

Appropriate feedback, provided in context, can improve teaching quality, however the teachers own interpretation of that data and how it is to be used plays a key role in understanding how changes to both delivery and content of learning may occur. It is known that many university lecturers do not consistently or systematically use feedback to improve their own teaching (Smith, 2008), and that their perceptions of students evaluations of their performance is generally negative (Aleamoni, 1999; Machingambi & Wadesango, 2011). The use of such feedback as part of a summative assessment of staff performance may also help to shape the attitudes of teaching staff towards this process. Making use of feedback to help develop teaching skills and ability is likely to result in an overall increase in such rating over time (Spooren & Mortelmans, 2006).

Self-reflection is one mechanism whereby teachers may be able to begin to evaluate their own performance and identify areas for improvement in practice and build upon aspects of their delivery of learning that are well received by learners. Regular student evaluation of performance can encourage reflective practice, and this in turn can lead to improvement in the delivery of teaching (Winchester & Winchester, 2011, 2014). In a recent paper Grant (2014) discusses the role of teachers as learners and suggests that the application of the methods for student assessment of teaching be applied to the teachers themselves as part of a constructionist paradigm (Grant, 2014). Self-assessment of teaching is not a new concept, as early as the 1970's researchers were proposing self-rating of teaching as a potential mechanism of improving performance (Centra, 1979). There is however a need to make a comparison between self-assessment and some form of external evaluation (i.e. Student feedback), and this will be most effective if teachers make their own assessment before obtaining such data (Airasian & Gullichson, 1994).

There has been some interest in exploring the relationship between teachers' self-assessment and student rating of teacher performance. Comparison of the ratings of 820 students on the performance of 52 instructors found only a modest relationship (correlation of 0.28, non –significant) between the staff and students appraisals of performance (Miron, 1988). Instructors tend to overestimate their own performance in comparison to student ratings (Shiloach & Fresko, 2002); interestingly in this study instructors were also asked to predict what ratings the students would provide in addition to their own self-assessment, with staff consistently predicating higher ratings than they received. However, other research has found that instructor predictions of student ratings closely match actual ratings, and that low rated instructors give higher estimates of student assessments than high rated instructors (Nasser & Fresko, 2006). A recent study comparing evaluation of teaching by students, teachers and management again found significant differences between staff and student ratings, and between management and

lecturer ratings (there was no difference between management and student ratings), leading the authors to conclude that self-assessment must be taken into consideration together with the student evaluation when evaluating staff performance (Taheri, Ryasi, Afshar, & Mofatteh, 2014).

A consistent finding of the research into the relationship between student feedback and instructor self-assessment of performance is that there is a lack of agreement between these ratings. Reasons for this discrepancy are less clear; it could be that instructors are not skilled at undertaking self-assessment, or that they may be unwilling to expose their weaknesses as they may be concerned that such information may be recorded and utilised in their own formal appraisals (Miron, 1988). In recent times it has become a requirement that university staff who are required to deliver teaching obtain a recognised educational qualification, and this may well involve a degree of reflective practice, therefore it may be that lecturers own educational background may be a factor that could influence their self-assessment of performance.

In our exploration of the relationship between staff self-assessment and student evaluation of teaching we have not determined any research that has been carried out on a post-graduate teaching programme. It will be of interest to determine if the findings for further and higher education (up to degree level) are replicated in a post-graduate / professional training course context. Therefore in this study we will explore the relationship between educator's self-assessment of performance and the assessment of the students themselves.

Methodology

All PsychD trainees (n=90) were asked to complete a weekly (anonymous) "Evaluation of teaching" for each formal lecture they attend in the period 2017/18. This form covered five areas of teaching performance, each assessed on a five point Likert scale:

1. How useful were the materials presented
2. How effective was the organisation of the session
3. To what extent were the learning objectives met
4. To what extent were service user and carer issues addressed
5. To what extent was difference and diversity covered

While completion of these forms is a course requirement, no record is kept of completion rates, and as such response rates vary from lecture to lecture. Completed forms are reviewed by the Research Director, and are then scanned and uploaded to a shared drive.

All staff (university and visiting) delivering teaching on the PsychD Clinical course in the year 2017/18 (n=45) were invited to participate in this research. A covering letter that outlined the rationale for the study was emailed to each lecturer, together with a link to an online survey (based on the trainees' evaluation of teaching form). Those lecturers that consented to participate were then instructed to complete the survey for each lecture they had delivered. Two weeks after the initial round of invitations, a reminder (with link to survey) was sent to all potential respondents to encourage participation.

Data collected was analysed using SPSS. Differences between the predicted and actual assessments scores on each item were calculated using Mann-Whiney U tests. Open ended questions asking about

lecturers views on how sessions could be improved were to be collated and analysed using content analysis.

Ethical approval

This study was submitted to the Faculty of Health & Medical Sciences Ethics Committee at the University of Surrey using the Fast Track process. A favourable ethical opinion (ref: FT-PSY-545-17) was given.

Results

We identified 45 lecturers (32 external, 13 university employees) who had delivered one or more teaching sessions to PsychD trainees in the year 2017/18 (September 2017 – April 2018). Following two rounds of invitations, data on 16 teaching sessions was collected (from a total of 60 delivered, 27% of the total) from 15 lecturers (33.3% response rate). Unfortunately at this time the feedback from 5 of those sessions was not available (currently under review), and therefore we were able to match the trainee and lecturer feedback for 11 sessions in total (18.3%). No respondents provided any answers to the open ended questions regarding ways in which sessions could be improved.

The number of completed trainee assessments varied from session to session, with a range of 15-31 completed questionnaires, with a median of 25 overall.

The following tables show the trainee / lecturer rating on each area of assessment for each session:

How useful were the materials presented?

Session	Lecturer rating	Trainee Rating (median)	Difference	P
1	3	5	-2	0.001
2	4	5	-1	0.004
3	4	5	-1	0.001
4	4	5	-1	0.001
5	3	5	-2	0.001
6	3	5	-2	0.001
7	4	5	-1	0.001
8	4	5	-1	0.001
9	4	4	0	0.013
10	3	4	-1	0.020
11	4	4	0	0.001

* Not significant

There were significant differences between ratings in every session (100%), with lecturers rating their performance on average one point lower than trainees.

How effective was the organisation of the session?

Session	Lecturer rating	Trainee Rating (median)	Difference	P
1	3	5	-2	0.001
2	4	4	0	0.011
3	5	4	+1	0.001
4	3	5	-2	0.001
5	3	4	-1	0.002
6	4	5	-1	0.001
7	4	4	0	0.020
8	4	5	-1	0.001
9	4	4	0	0.002
10	3	4	-1	0.026
11	4	4	0	0.605*

* Not significant

There were significant differences between ratings in almost every session (91%), with lecturers rating their performance on average one point lower than trainees.

To what extent were the learning objectives met?

Session	Lecturer rating	Trainee Rating (median)	Difference	P
1	3	5	-2	0.001
2	4	5	-1	0.001
3	4	5	-1	0.001
4	4	5	-1	0.001
5	3	5	-2	0.001
6	4	5	-1	0.001
7	5	5	0	0.001
8	4	5	-1	0.001
9	4	4	0	0.008
10	3	4	-1	0.002
11	4	4	0	0.851*

* Not significant

There were significant differences between ratings in almost every session (91%), with lecturers rating their performance on average one point lower than trainees.

To what extent were service user and carer issues addressed?

Session	Lecturer rating	Trainee Rating (median)	Difference	P
1	1	4.5	-3.5	0.026
2	4	5	-1	0.001
3	3	5	-2	0.001
4	2	5	-3	0.001
5	2	5	-3	0.001
6	4	5	-1	0.001
7	3	5	-2	0.001
8	4	4	0	0.197*
9	4	4	0	0.414*
10	5	5	0	0.180*
11	2	5	-3	0.001

* Not significant

There were significant differences between ratings in most sessions (73%), with lecturers rating their performance on average one and a half points lower than trainees.

To what extent was difference and diversity covered?

Session	Lecturer rating	Trainee Rating (median)	Difference	P
1	1	5	-4	0.016
2	3	5	-2	0.001
3	2	4	-2	0.001
4	2	4	-2	0.001
5	2	4	-2	0.001
6	5	5	0	0.317*
7	3	4	-1	0.001
8	2	3	-1	0.001
9	2	4	-2	0.001
10	5	5	0	0.157*
11	3	5	-2	0.001

* Not significant

There were significant differences between ratings in almost every session (82%), with lecturers rating their performance on average two points lower than trainees.

Discussion

One third of all lecturers on the PsychD programme responded to the survey, providing self-assessment of their teaching performance on approximately one quarter of all delivered sessions. We were able to match trainee and lecturer assessments on 73% of the submitted data, representing just under one fifth of all teaching sessions during the study period.

The data show significant differences in trainee and lecturer assessment of teaching performance in all areas of assessment across almost every session, with lecturers consistently underestimating their performance compared to the trainee evaluations of the same session.

Response rates to this survey were low (33%) therefore it is likely that our sample is not representative of the broader teaching staff, however we have been able to assess teaching on 20% of the programme across the study period, and so our findings may be indicative of a broader trend, and certainly reveal an issue that may well be of interest to course providers.

Our observation that teaching staff consistently rate their performance lower that reported by trainees suggests that there is an opportunity to provide positive feedback. While staff do receive copies of the trainee feedback (on average 8 weeks after they delivered the teaching), at present they are not invited to reflect upon their own assessment of their performance or to consider how their assessment may differ from that of the trainees. I believe that such a reflection would afford a teachable moment – an opportunity to bolster self confidence and consequently to further improve upon teaching performance. As such the recommendation that teaching staff assess their own performance (at the time of delivery) is offered to course convenors, and that mechanisms whereby that self assessment can be compared to assessment by trainees are developed.

While we have observed that in the majority of cases teaching staff has underestimated their performance, there are (limited) examples of where performance has been over-estimated. I suggest that in these cases, provision of feedback would also be of benefit – in these cases reflection could result in acknowledgement of the need for further training and support.

This study is the first to explore concordance between lecturer and trainee assessments of teaching performance in the context of a professional training programme, finding that unlike undergraduate programmes staff have a tendency to under-estimate their performance. There are two clear distinctions between professional training programmes and undergraduate degree programs; firstly the majority of teaching is delivered by professionals working in the field, who are not required to undertake any formal teacher training, and secondly the trainees themselves are postgraduates with at least one years experience of working in the field. These differences might explain, in part, our observations, however this study is limited by its sample selection and low response rate. None the less, it is clear that teaching staff of the professional doctorate programme do tend to rate their teaching performance lower than the trainees. While the effect that this may have on future teaching or retention of teaching staff is unclear, it would be interesting to explore this finding further using a more robust sampling frame and methodology. It is unfortunate that no teaching staff provided any suggestions as to how their performance might be improved, and this is another area where future

research might consider a different approach to data collection, perhaps face to face interviews or focus groups might encourage participants to engage with these issues. Despite the limitations of this study, a clear message to teaching staff on the professional doctorate programme is evident: You're better than you think you are.

References

Airasian, P., & Gullichson, A. (1994). Examination of teacher self-assessment. *Journal of Personal Evaluation in Education, 8*(2), 195-203.

Aleamoni, L. (1999). Student rating myths versus reserach facts from 1924 to 1998. *Journal of Personnel Evaluation in Education, 13*(2), 153-166.

Centra, J. (1979). *Determining Faculty Effectiveness. Assessing Teaching, Research and Service for Personnel Decisions and Improvement.* San Fancisco: Josey-Bass.

Cohen, P. (1980). Effectiveness of student-rating feedback for improving college instruction: A meta analysis of findings. *Research in Higher Education, 13*(4), 321-341.

Grant, K. (2014). Comparing Student and Teacher Self-Assessment Practices. In P. P (Ed.), *Proceedings of the IDEAS: Rising to Challenges conference*: University of Calgary.

Kember, D., Leung, D., & Kwan, K. (2002). Does the use of Student Feedback questionnaires improve the overall quality of teaching? *Assessment & Evaluation in Higher Education, 27*(5), 411-425.

Lang, J., & Kersting, M. (2007). Regular feedback from student ratings of instruction: Do college teachers improve their ratings in the long term? *Instructional Science, 35*(3), 187-205.

Lizzo, A., Wilson, K., & Simons, R. (2002). University students perceptions of the learning environment and academic outcomes: implications for theory and practice. *Studies in Higher Education, 27*(1), 27-52.

Machingambi, S., & Wadesango, N. (2011). University lecturers perceptions of students evaluations of their institutional practices. *Anthropologist, 13*(3), 167-174.

Marsh, H. (1987). Students evaluations of university teaching: Research findings, methofdological issues and directions for future research. *International Journal of Edcational Research, 11*(3), 253-388.

Miron, M. (1988). Students' evaluation and instructors self evaluation of university instruction. *Higher Education, 17*(2), 175-181.

Nasser, F., & Fresko, B. (2006). Predicting student ratings: The relationship between actual student ratings and instructor predictions. *Assessment and Evaluation in Higher Education, 31*(1), 1-18.

Ovando, M. (1994). Constructive feedback: A key to successful teaching and learning. *International Journal of Educational Management, 8*(6), 19-22.

Seldin, P. (1995). *Improving college teaching*. Bolton, MA: Anker.

Seldin, P. (1997). Using student feedback to improve teaching. *New Directions for Teaching and Learning, 19*(37), 89-97.

Shiloach, Y., & Fresko, B. (2002). Instructor educators and students evaluate college instructiojn: do they agree? In B. Fresko & D. Kfif (Eds.), *An on-going dialogue: instructor education and educational practice*. Tel Aviv: Mofet Institute.

Smith, C. (2008). Building effectiveness in teaching through targeted evaluation and response: Connecting evaluation to teaching improvement in higher education. *Assessment & Evaluation in Higher Education, 33*(5), 517-533.

Spooren, P., & Mortelmans, D. (2006). Teacher professionalism and student evaluation of teaching: will better teachers recieve higher ratings, and will better students give higher ratings? *Educational Studies, 32*(2), 201-214.

Taheri, M., Ryasi, H., Afshar, M., & Mofatteh, M. (2014). Comparison between student rating, faculty self-rating and evaluation of faculty members by heads of respective academic departments in the School of Medicine in Birjand University of Medical Sciences in Iran. *Journal of Education and Health Promotion, 3*(34).

Winchester, T., & Winchester, M. (2011). Exploring the impact of faculty reflection on weekly student evaluations of teaching. *International Journal of Personnel Evaluation in Education, 14*(2), 179-192.

Winchester, T., & Winchester, M. (2014). A longitudional investigation of the impact of faculty reflective practices on students' evaluations of teaching. *British Journal of Educational Technology, 45*(1), 112-124.