

**Use of the ‘Stop, Start, Continue’ method is associated with the production of constructive qualitative feedback by students in higher education**

**Alice E. Hoon<sup>1</sup>, Emily J. Oliver<sup>2</sup>, Kasia Szpakowska<sup>3</sup> and Philip M. Newton<sup>1\*</sup>**

<sup>1</sup> College of Medicine, Swansea University, Swansea, Wales, United Kingdom, SA2 8PP tel 01792 513400

<sup>2</sup> School of Applied Social Sciences, Durham University tel 0191 33 46997

<sup>3</sup> College of Arts and Humanities, Swansea University, Swansea, Wales, United Kingdom, SA2 8PP tel 01792 513104

\* Author to who correspondence should be addressed – [p.newton@swansea.ac.uk](mailto:p.newton@swansea.ac.uk)

**Abstract** Students in higher education are increasingly asked to give feedback on their education experience, reflecting an increase in the importance attached to that feedback. Existing literature demonstrates that qualitative student feedback is valued and important, yet there has been limited evaluation of the means by which qualitative student feedback is collected, and how collection method influences the depth of feedback from students. We evaluated the depth of feedback written by students on programmes at three different universities in the UK, using an established evaluation instrument. We found that the use of a structured proforma ('Stop, Start, Continue') was associated with feedback of greater depth than that produced using free-text entry. We then evaluated the effect of switching from a free-text to a structured proforma at one institution and found, again, that the structured proforma was associated with feedback of greater depth. In addition, students indicated a preference for the structured proforma.

**Keywords:** student feedback, Stop Start Continue, proforma, student satisfaction

## Introduction

Many studies have evaluated the provision of qualitative written feedback to students, and the impact of strategies aimed at improving that feedback. In a synthesis of this literature, Valerie Shute (Shute, 2008) described the existence of multiple types of feedback (e.g. 'verification', 'error flagging') and observations that feedback utility is influenced by many interacting parameters, such as the timing of feedback delivery (e.g. immediate versus delayed), the level and prior achievement of the learner, the subject(s) being studied, the medium used to deliver the feedback (e.g. written, oral, face to face), and the complexity of the task. This literature has informed general agreement on best practice guidelines for the production and use of feedback written to students, such as ensuring that the feedback is clear, specific, objective and focused on the task (not the learner). Despite the abundant writing in this area, a review of the literature on feedback to students, written in 2003 (Mutch 2003) concluded that the principles of good feedback were under-researched and many of the studies conducted had a limited sample size.

The qualitative feedback written by students to their educators has received much less attention in the education research literature. It seems reasonable to assume that some of the principles described by Shute could be applied to the generation of feedback by students for teachers and their institutions, but there is a lack of evidence to inform policy and practice in this area. There are some obvious important differences between feedback written by students and feedback written by teachers, most notably in the fundamental purposes of the feedback, power differentials and hierarchies, and the timing and collection methods for that feedback.

Despite the relative paucity of evidence to support its use, the impact of student-generated qualitative feedback has gradually escalated. The use of student-generated feedback to improve teaching is well-established, although questions consistently arise about its effectiveness for that purpose, and the means by which it is gathered (e.g. see Berk 2012 for a recent review). More recent developments in the United Kingdom include the National Student Survey (NSS), used to collect data (largely quantitative) about the student experience and used in the evaluation of higher education institutions (Buckley 2012). Similarly, student evaluation of teaching (SET) forms a significant part of overall performance evaluations of teachers in the USA. Many years of research have resulted in a number of validated, reliable instruments for SET, again mostly quantitative, and there is good evidence that one effect of SET is to improve future teaching (Marsh and Roche 1997). However, unresolved controversies persist around a number of confounding variables which can influence evaluations, such as a teacher's 'attractiveness', gender, race/ethnicity and age, as well as factors related to the course such as predicted/expected grades and the perceived difficulty of the course (Basow and Martin, 2012).

The use of the aforementioned quantitative measures can provide important information to different stakeholder groups, and the NSS is one of the latest in a series of formal instruments used to collect quantitative data. For course development and teaching improvement there is a need to collect and analyse qualitative feedback provided by students (Grebennikov and Shah 2013). Students value free text comments in feedback (Kindred and Mohammed 2005), and qualitative data also add context and detail to issues which arising in the quantitative data

(Symons 2006). However, the resources required to conduct a meaningful analysis of qualitative data are much greater than for quantitative feedback, which is often a deterrent to the collection of open-ended qualitative feedback (Richardson 2005). Hence there is a need to ensure that the collection of qualitative data is maximally efficient, not directly replicating qualitative feedback while retaining opportunities for full expression. For example it has recently been shown that the use of a structured feedback proforma is, when compared to a free text proforma, associated with the production of more abundant and deeper qualitative feedback written to students by the markers of written assessments at a UK medical school (Newton, Wallace, and McKimm 2012).

Here we applied a similar methodology to the analysis of instructor feedback to that written by students on their learning experience. Three programmes at universities in the UK with pre-existing variation in their feedback collection methodologies were purposively sampled. Two programmes used a structured format, similar to the ‘Stop, Start, Continue’ model. This is a well established method of collecting feedback in many disciplines, including education (e.g. see George and Cowan 1999, 28). It asks structured questions that collect student views on what educators should no longer do (stop), should consider adding (start), and features that should be retained (continue). The third programme collected feedback using a blank, free-text only proforma.

We found that the use of a structured, questioning proforma was associated with the production of deeper feedback by students. Similarly, a switch from a blank to a structured proforma was associated with an increase in the depth of feedback produced, and a positive evaluation of the new form by students.

## **Method**

Written student feedback from three different courses at two different institutions in the United Kingdom was evaluated. This module-level feedback had been collected as part of the routine evaluation of each programme.

Course A was a graduate-entry medical degree, thus an undergraduate course undertaken by students who already have an undergraduate degree. Each intake has ~70 students. Teaching in the first two years of the course (on which the current study is focused) is arranged around discrete teaching weeks. Students provide qualitative feedback weekly on each university-based teaching week, using a proforma that has a free-text entry box for every teaching session occurring that week.

Course B was an undergraduate Bachelor of Science degree, with an annual intake of between 80 and 100 students. Teaching occurs in modules taught over 12 or 24 weeks. Students provide qualitative feedback on their course at the end of each module, and at the end of each academic year. Feedback analysed from Course B was collected at the end of twelve different modules using the ‘Stop Start Continue’ format. This format consists of presenting three text boxes containing the words Stop, Start, and Continue. In the Stop box

students are given the opportunity to make suggestions about what they do not find useful or helpful, for example, ‘stop speaking so quickly’. In the Start box students have the opportunity to make suggestions or give ideas about what they think could be started which may aid their learning, for example, ‘start giving out class handouts’. The Continue box allows students to comment on best practice and aspects of teaching that they find useful, for example, ‘continue the use of class quizzes’.

Course C was a level 2 undergraduate module in the Arts and Humanities, taken by students who have already completed a module in the same subject area. Each intake has 25-40 students. Teaching occurs in the module over 11 weeks in the first semester. Students provide qualitative feedback at the end of each module. Feedback analysed for Course C was collected at the end of four modules. Students were given a feedback form containing text box for each of the following questions: 1) What was the best thing about the module? 2) What would you change about the module? 3) Any other comments?

### **Coding of feedback**

Qualitative feedback was coded using a feedback system similar to those developed by Newton, Wallace, and McKimm (2012) and Brown and Glover (2006). Every statement in the feedback was coded according to two criteria – the first being whether its overall theme was that the issue being commented on was ‘positive’, ‘negative’, or ‘adequate’ (coded as P, N, or A, respectively), and the second being the depth of the feedback. For example if a statement merely indicated at the most basic level whether something was good or not good, it was coded as 1 (Descriptive); if a statement indicated why something was or was not good it was coded as 2 (Qualified); and if a statement included constructive suggestions for change or development it was categorised as a 3 (Constructive). Statements made by a student that did not relate to the content of the lecture/module in question, for example reporting that they did not attend a lecture, were classed as ‘Other’. See Table 1 for more information.

**Table 1. Classification of feedback statements**

| <b>Feedback category</b> | <b>Basic theme of statement</b>                         |
|--------------------------|---------------------------------------------------------|
| <b>P1</b>                | This was good.                                          |
| <b>P2</b>                | This was good because....                               |
| <b>P3</b>                | This was good because... but would be even better if... |
| <b>N1</b>                | This was bad...                                         |
| <b>N2</b>                | This was bad because....                                |
| <b>N3</b>                | This was bad because... and would be better if...       |
| <b>A1</b>                | This was adequate...                                    |
| <b>A2</b>                | This was adequate because...                            |
| <b>A3</b>                | This was adequate because... and would be better if...  |

**Other** e.g. I did not attend, this lecture was cancelled etc.

(coding adapted from Newton, Wallace, & McKimm, 2012)

Although all feedback was labelled as positive, negative or adequate, the authors were most interested in the quality of the feedback, that is, how useful or constructive each statement was, rather than how well the lecture or module had been received by the students. To ensure consistency, all coding was performed by only one author (AH) after training by another (PN).

## Results and discussion

Students in Course A made a total of 1072 statements, students in Course B made a total of 264 statements, and students in Course C made 131 statements. The detailed coding results for each depth category are shown by percentage in Table 2, with a representation of the feedback collapsed by category shown in Figure 1. Percentages were used to factor out any effect of cohort size.

**Table 2 Percentage of statements in each feedback category by course.**

|                                       | Depth           |     |     |               |     |     |                 |      |     |       |
|---------------------------------------|-----------------|-----|-----|---------------|-----|-----|-----------------|------|-----|-------|
|                                       | 1 (Descriptive) |     |     | 2 (Qualified) |     |     | 3(Constructive) |      |     | Other |
|                                       | P               | N   | A   | P             | N   | A   | P               | N    | A   |       |
| <b>Course A (Blank)</b>               | 42.2            | 0.4 | 2.2 | 33.5          | 4.2 | 2.1 | 4.3             | 9.1  | 0.8 | 1.4   |
| <b>Course B (Stop/Start/Continue)</b> | 2.7             | 2.3 | -   | 3.4           | 0.1 | 0.8 | 35.6            | 54.6 | -   | -     |
| <b>Course C (Best/Change/Other)</b>   | 2.3             | -   | -   | 61.8          | 1.5 | -   | 1.5             | 29.8 | -   | 3.1   |
| <b>Total</b>                          | 47.1            | 2.6 | 2.2 | 98.7          | 5.8 | 2.8 | 41.4            | 93.4 | 0.8 | 4.4   |
| <b>Mean</b>                           | 15.7            | 0.9 | 0.7 | 32.9          | 1.9 | 0.9 | 13.8            | 31.1 | 0.3 | 1.5   |
| <b>SD</b>                             | 22.9            | 1.2 | 1.2 | 29.2          | 2.1 | 1.0 | 18.9            | 22.8 | 0.5 | 1.5   |

The data represented in Table 2 show that the feedback collected from Course A, using a blank proforma, were largely depth category 1 and 2 responses, whereas very few category 1 responses were made by respondents in Course B and C which used more structured approaches. In contrast the feedback collected from Course B was largely in the most constructive category 3. Thus although more feedback was collected from Course A, the depth of that feedback was lower than in the two other courses, and the descriptive nature of the comments were of a nature that could be simply captured using a quantitative method.

(Figure 1 should go here)

**Figure 1. Percentage of statements in each feedback category by course and the type of proforma used to collect the feedback.**

As there was much variation across the three courses in the percentage of statements attributed to each feedback category, a further analysis was conducted in which the percentage of statements for each question type was analysed. This was conducted on the feedback collected for Course B and Course C as the feedback format asked specific questions. This analysis could not be conducted for Course A as students were only asked to comment on each teaching session. Table 3 shows that asking students what they would like to stop occurring and what they would like to start occurring both tended to result in N3-type responses. Asking students what think should be continued typically resulted in P3-type responses. For Course C, it was found that asking students “what did you like about this module” typically resulted in P2 responses, whereas “what would you change about this module” generally resulted in N3 responses.

**Table 3. Percentage of statements in each feedback category by question type.**

|                 |                               | Depth           |    |    |      |               |     |      |                  |    |       |
|-----------------|-------------------------------|-----------------|----|----|------|---------------|-----|------|------------------|----|-------|
|                 |                               | 1 (Descriptive) |    |    |      | 2 (Qualified) |     |      | 3 (Constructive) |    |       |
|                 |                               | P1              | N1 | A1 | P2   | N2            | A2  | P3   | N3               | A3 | Other |
| <b>Course B</b> | <b>Stop</b>                   | 0.7             | -  | -  | -    | 0.7           | -   | -    | 20.9             | -  | -     |
|                 | <b>Start</b>                  | -               | -  | -  | -    | -             | -   | -    | 36.2             | -  | -     |
|                 | <b>Continue</b>               | 1.5             | -  | -  | 3.7  | -             | 0.8 | 34.7 | 1.5              | -  | -     |
| <b>Course C</b> | <b>Best thing<sup>1</sup></b> | 1.6             | -  | -  | 48.8 | -             | -   | -    | -                | -  | -     |
|                 | <b>Change<sup>2</sup></b>     | 0.8             | -  | -  | 1.6  | 1.6           | -   | 1.6  | 29.5             | -  | -     |
|                 | <b>Comments<sup>3</sup></b>   | -               | -  | -  | 11.6 | -             | -   | -    | 0.8              | -  | 2.3   |
| <b>Total</b>    |                               | 4.6             | -  | -  | 65.8 | 2.3           | -   | 36.3 | 88.8             | -  | 2.3   |
| <b>Mean</b>     |                               | 0.8             | -  | -  | 11.0 | 0.4           | 0.8 | 7.3  | 14.8             | -  | 0.4   |
| <b>SD</b>       |                               | 0.7             | -  | -  | 19.1 | 0.6           | -   | 15.4 | 16.1             | -  | 1.0   |

<sup>1</sup>What was the best thing about this module?

<sup>2</sup>What would you change about this module?

<sup>3</sup>Any other comments?

Figure 2 below indicates that for Course B, the Stop/Start/Continue format resulted in students giving feedback which made constructive suggestions for change. For Course C, asking students what they liked best about the course typically resulted in responses regarding what was good about the course and why, whereas specifically asking what students would change about the course, resulted in students making Category 3 suggestions for improvements. When students were asked if they had any other comments, the responses

given tended to be Category 2 statements which, although useful and qualified, did not offer constructive suggestions for change or development.

Figure 2 should go here

**Figure 2. Percentage of statements in each feedback category by question type.**

Taken together, these findings suggest that the specificity of questions asked in evaluation forms may influence the constructiveness of the feedback obtained. Using only open-ended questions, such as the format used on Course A, was associated with feedback which was mostly descriptive, and may therefore be limited in terms of enabling course development. However, when questions were phrased to specifically ask what aspects of teaching should be changed or continued as per Course B, much more of the feedback collected was of a type which made suggestions that could potentially inform constructive change in teaching practice.

One interpretation of these findings may be that more constructive qualitative feedback can be obtained when using a structured form. However this interpretation is subject to many confounds - the feedback was collected from different students on three different courses, in different disciplines, at two different institutions, with students at different stages (undergraduate vs postgraduate) and at different frequencies (weekly vs at the end of a semester). Thus it is difficult to say with any certainty whether the differences in the quality of the feedback were a direct function of the feedback format or due to another variable. To understand this further, we designed a second study.

## **Study 2**

In this study, the qualitative feedback form for Course A was changed from the open ended format used previously, to a modified Stop/Start/Continue format. We applied the modified format over four teaching weeks across two cohorts, and feedback collected was then evaluated as in Study 1. Given our evidence that the open-ended form was associated with poorer quality feedback, concurrent use of this in control courses or institutions was not appropriate. The analysis was thus compared with the feedback that had been completed for the same teaching weeks in the previous academic year, to assess whether changing the format of feedback was associated with better quality feedback.

## **Method**

At the end of each teaching week, students were presented with a feedback form which asked the following questions:

- 1) Stop: What would you like us to stop doing?
- 2) Start: What suggestions do you have?
- 3) Continue: What is being done well that you would like to continue?
- 4) Good practice/highlights?



- 5) Any other comments?
- 6) Any comments about the new feedback form

Questions 4 was included as much of the teaching on Course A was given by external speakers who may value course feedback as evidence of their own professional development and engagement with education. Question 5 was included to allow students to make more general comments relating to the course that may not have fitted within the Stop/Start/Continue structure, in keeping with literature suggesting that open-ended questions are valued by students and may result in important feedback (Grebennikov and Shah 2013; Kabanoff, Richardson, and Brown 2003).

The Stop/Start/Continue form was distributed at the end of four teaching weeks to the first and second year cohorts, therefore representing eight teaching weeks in total. All feedback was analysed and coded as in Study 1. Feedback for the same teaching weeks in the previous academic year of Course A were obtained and coded in the same way. These weeks were different to the weeks analysed in Study 1.

## Results

In the original free text format, the Year 1 students made a total of 872 statements and the Year 2 students made a total of 611 statements. When presented with the Start Stop Continue format, the Year 1 students made a total of 313 statements, whereas Year 2 made a total of 204 statements. Results from the coding of these statements are shown in Table 4, which shows that, in agreement with the findings from Study 1, the original unstructured feedback format yielded a high number of Category 1 responses, with 44.52% of statements falling within the P1 category.

**Table 4. Percentage of statements in each feedback category in the original free text format and the Stop/Start/Continue format.**

|                            | Depth           |      |      |               |       |      |                  |       |   |       |
|----------------------------|-----------------|------|------|---------------|-------|------|------------------|-------|---|-------|
|                            | 1 (Descriptive) |      |      | 2 (Qualified) |       |      | 3 (Constructive) |       |   | Other |
|                            | P               | N    | A    | P             | N     | A    | P                | N     | A |       |
| <b>Free text only</b>      | 44.52           | 1.31 | 2.62 | 23.73         | 15.55 | 0.98 | 3.27             | 6.06  | 0 | 1.96  |
| <b>Stop/Start/Continue</b> | 3.92            | 0    | 0    | 28.43         | 12.75 | 0    | 25.98            | 24.02 | 0 | 4.9   |

We again collapsed our analysis into depth categories only. For the Stop/Start/Continue format however, 50% of statements were classed as Category 3 (see **Figure 3**). With regard to the number of statements classed as Category 2, there was little change between the original format and the Start Stop Continue format.

(Figure 3 should go here)

**Figure 3. Percentage of statements in each feedback category in the original free text feedback format and the Stop/Start/Continue format.**

### Analysis by question

To identify whether a particular feedback question was more likely to result in a particular type of statement, an analysis by question type was conducted. Table 5 shows that ‘Stop’ almost always resulted in N3-type statements, ‘Start’ typically resulted in N3 or P3 statements, whereas Continue tended to result in P3 statements. The majority of P2 responses were written in the Good practice section. There were no N1 or category A statements.

**Table 5. Percentage of statements for each feedback category per question.** No ‘A’ category statements were found

|                      | <b>P1</b> | <b>P2</b> | <b>P3</b> | <b>N1</b> | <b>N2</b> | <b>N3</b> | <b>Other</b> | <b>Total</b> |
|----------------------|-----------|-----------|-----------|-----------|-----------|-----------|--------------|--------------|
| <b>Stop</b>          | -         | 0.2       | 0.2       | -         | 1.9       | 15.4      | 0.9          | 18.6         |
| <b>Start</b>         | 0.2       | 3.2       | 6.2       | -         | 3.4       | 10.9      | -            | 23.9         |
| <b>Continue</b>      | 0.9       | 6.8       | 15.2      | -         | 0.2       | 0.4       | -            | 23.5         |
| <b>Good practice</b> | 1.3       | 15.4      | 7.7       | -         | 1.1       | 0.6       | 0.2          | 26.3         |
| <b>Comments</b>      | 0.9       | 2.6       | 0.2       | -         | 1.1       | 2.1       | 0.6          | 7.5          |
| <b>Total</b>         | 3.2       | 28.1      | 29.4      | -         | 7.7       | 29.4      | 1.7          |              |

### Student evaluation of the new form design

Students were asked to comment on the new Stop/Start/Continue design, using an open-ended question. The depth and volume of feedback obtained in response to this question were limited, and the only theme emerging was a preference (or not) for the new form. Of the 31 comments, received, 20 were positive and 11 were negative, meaning that 65% of student comments indicated a preference for the new form. Within the positive theme, individual comments indicated that students found the new form “better”, “more constructive” and observed that it allowed them to give feedback on the week as a whole.

### General discussion

Study 1 reported that the format of a feedback form may have implications for the depth of qualitative feedback that is obtained. Specifically, the use of forms which give the opportunity to provide open, free text responses, was associated with large quantities of descriptive feedback that did not necessarily allow for constructive change or course development. The Stop/Start/Continue format used in Course B, which specifically asked the

responder to make suggestions for change, was associated with less feedback in terms of quantity of statements, but with highly constructive feedback in those statements given. A similar pattern was found for Course C whereby asking students what they would change about a course was more likely to result in constructive comments, whereas asking what they liked most about a course tended to result in more descriptive statements about the content of the course.

The findings of Study 1 suggest that the structure of a feedback form might have a significant impact upon the type of qualitative feedback that is collected, and that the specific questions asked may influence the responses given. A key limitation of this study, however, was that it compared the feedback of three different groups of students on three very different courses, at two different institutions, with feedback collected at different times and with different frequencies. Study 2 addressed this limitation by using a cohort design with one module, demonstrating that the Stop/Start/Continue format increased use of constructive statements relative to the open-ended format. Additionally, Study 2 found that whilst students appeared to make fewer statements using the Start Stop Continue format, the quality of feedback improved. This is an important finding as there is an ever-increasing expectation on students to provide regular course feedback, and qualitative data are a rich source of important information that may aid course development (Grebennikov and Shah 2013; Symons 2006), yet their utility is constrained by the resources required to analyse them (Richardson 2005). Thus it is important to identify ways in which the richness of qualitative feedback can be retained (and even enhanced) while reducing the resource load for analysis. Lastly, it was apparent that students, as a group, preferred the Stop/Start/Continue format. Having students engaged with providing feedback is a fundamentally important part of obtaining successful feedback (Richardson 2005; Buckley 2012), and thus the preferences expressed here for the Stop/Start/Continue format further support its use.

There are some additional limitations which should be considered in the interpretation of the results presented here. The tool we have used to code the feedback assigned categories based upon a simple depth criterion. Obviously such an approach will not account for all the complexity of qualitative feedback collected from students. For example, one N2-type statement which points out a significant limitation (e.g. ‘the content of lecture X was factually inaccurate’) may be of much greater importance than an N3 statement on a much less significant topic (e.g. ‘the handouts for lecture Z were useful but having page numbers would have made them easier to use’). However, having analysed feedback from across multiple teaching episodes for each course type, the impact of this limitation should be reduced.

Another limitation is that the tool used (Newton, Wallace, and McKimm, 2012; Brown and Glover, 2006) was developed for the analysis of feedback written to, rather than by, students, with a focus on constructively improving student achievement. However much of the impetus for collecting student feedback is to inform course development on the basis of student views on current teaching effectiveness (Berk 2012) and the relative usefulness of category 1 versus category 3 statements seems applicable to that end. Nevertheless, future developments of this work might involve the development of specific instruments for the coding of student-

generated feedback, that take account of all the motivations for collecting student feedback, such as teacher appraisal, benchmarking, quality assurance and review (Buckley 2012). Future work may also wish to elicit teaching staff's perceptions of the utility of the different types of feedback provided by students, as they are one of the key end users of the feedback.

In summary, much of the literature in assessment and feedback has focused on the generation of feedback for students, and on quantitative feedback by students, with far less attention paid to the means by which qualitative feedback written by students can be optimised for course development and other outcomes. Although there have been developments in this area (e.g. see Buckley 2012), and the use of student feedback for appraisal of teachers is well established in the USA and other countries (Basow and Martin 2012; Berk 2012), there was a need for further study to ensure that the maximal benefit is harvested from the effort put into the generation of feedback by students. The present findings highlight the importance of using feedback forms which are specifically tailored towards the type of evaluation that is required, as the specific questions asked on a feedback form impact upon the type of feedback that is collected, and therefore the extent to which that feedback may be useful. These findings suggest that the purpose of an evaluation may be key to determining the most appropriate format by which feedback is collected. If the purpose of an evaluation is to identify how a programme can be improved for example, it is important to include questions that explicitly ask what improvements can be made. If, however, the purpose of evaluation is to identify more general views and opinions of a respondent, then open questions may be appropriate. Whilst the use of open, free text questions such as those used by Course A in Study 1 may be a source of useful information, such formats do not lend themselves to collecting feedback which could inform development, and the volume of feedback generated may provide an impediment to its use. It is likely that a combination of both structured questions and free text may be preferable, as this allows students to make suggestions for change whilst also ensuring they have the opportunity to express any other views that may not strictly fit within the Stop, Start or Continue headings.

The findings reported here support the use of a modified Stop/Start/Continue format which allows for the efficient collection of qualitative student feedback while retaining opportunities for full expression.

### **Short Biographical Note for each author**

**Dr Alice Hoon** is a Lecturer in Assessment and Evaluation at the College of Medicine at Swansea University. She contributed to study design and manuscript preparation. She also coded and analysed all of the feedback described in the study.

**Dr Emily Oliver** is a Lecturer in Sport and Exercise Psychology in the School of Applied Social Sciences at Durham University. Her research interests include motivation and engagement. She contributed by facilitating data collection and editing the manuscript.

**Dr Kasia Szpakowska** is an Associate Professor in the College of Arts and Humanities at Swansea University. She contributed by facilitating data collection.

**Dr Philip Newton** is an Associate Professor in the College of Medicine at Swansea University. His research interests include evidence-based education. He contributed to study design, supervision and manuscript preparation.

**Disclosure statement** – the authors have no financial or other interests/benefits, from the application of the findings of this evaluation.

## References

- Basow, Susan A, and Julie L Martin. 2012. 'Bias in Student Evaluations'. In *Effective Evaluation of Teaching*, 40–49. Society for the Teaching of Psychology.  
<http://teachpsych.org/ebooks/evals2012/index.php>.
- Berk, Ronald A. 2012. 'Top Five Flashpoints in the Assessment of Teaching Effectiveness'. *Medical Teacher*, October, 1–12. doi:10.3109/0142159X.2012.732247.
- Brown, Evelyn, and Chris Glover. 2006. 'Research Article: Written Feedback for Students: Too Much, Too Detailed or Too Incomprehensible to Be Effective?' May 26.  
<http://www.bioscience.heacademy.ac.uk/journal/vol7/beej-7-3.aspx>.
- Buckley, Alex. 2012. 'Making It Count'. UK Higher Education Academy resources.  
[http://www.heacademy.ac.uk/resources/detail/nss/Making\\_it\\_count](http://www.heacademy.ac.uk/resources/detail/nss/Making_it_count).
- George, Judith W., and John Cowan. 1999. *A Handbook of Techniques for Formative Evaluation: Mapping the Student's Learning Experience*. Psychology Press.
- Grebennikov, Leonid, and Mahsood Shah. 2013. 'Student Voice: Using Qualitative Feedback from Students to Enhance Their University Experience'. *Teaching in Higher Education* 18 (6): 606–18. doi:10.1080/13562517.2013.774353.
- Kabanoff, B, A Richardson, and S Brown. 2003. 'Business Graduates' Perceptions of the Quality of Their Course: A View from Their Workplace'. *Journal of Institutional Research* 12 (2): 1.
- Kindred, Jeannette, and Shaheed N. Mohammed. 2005. "'He Will Crush You Like an Academic Ninja!": Exploring Teacher Ratings on Ratemyprofessors.com'. *Journal of Computer-Mediated Communication* 10 (3): 00–00. doi:10.1111/j.1083-6101.2005.tb00257.x.
- Marsh, Herbert W., and Lawrence A. Roche. 1997. 'Making Students' Evaluations of Teaching Effectiveness Effective: The Critical Issues of Validity, Bias, and Utility'. *American Psychologist* 52 (11): 1187–97. doi:10.1037/0003-066X.52.11.1187.
- Mutch, Alistair. 2003. 'Exploring the Practice of Feedback to Students'. *Active Learning in Higher Education* 4 (1): 24–38. doi:10.1177/1469787403004001003.
- Newton, Philip M, Melisa J Wallace, and Judy McKimm. 2012. 'Improved Quality and Quantity of Written Feedback Is Associated with a Structured Feedback Proforma'. *Journal of Educational Evaluation for Health Professions* 9: 10. doi:10.3352/jeehp.2012.9.10.
- Richardson, J. 2005. 'Instruments for Obtaining Student Feedback: A Review of the Literature'. *Assessment & Evaluation in Higher Education* 30 (4): 387–415.  
doi:10.1080/02602930500099193.
- Shute, Valerie J. 2008. 'Focus on Formative Feedback'. *Review of Educational Research* 78 (1): 153–89. doi:10.3102/0034654307313795.
- Symons, R. 2006. 'In Their Own Words: Finding out What Students Think about Their University Learning Experience'. <http://www.itl.usyd.edu.au/synergy/article.cfm?articleID=281>.