A discipline specific factor analysis approach to using student surveys for improvement

K. Patrick, A. Bedford, S. Romagnano, M. Bedford, and J. Barber. RMIT University, Melbourne, Australia.

Correspondence to:
Associate Professor Kate Patrick
Quality Unit, RMIT University, GPO Box 2476V, Melbourne, Victoria 3001, Australia.
+61 3 9925 2641.
kate.patrick@rmit.edu.au.

ABSTRACT

Like other universities, RMIT recognises the significance of graduates’ ratings of their experience and has had a long-term commitment to improving student learning. As at other universities, RMIT’s standard subject-level survey (the Course Experience Survey) incorporates items from the national Course Experience Questionnaire, with the aim of eliciting student views and prompting improvements which will take effect before the students graduate. The university has been seeking strategies to make the results of these surveys more accessible to academic staff, so that staff can use them as a starting point for change.

The current project is part of this work. The starting point is discipline based analysis of the university’s CES data. Surveys were stratified into fifty disciplines, and categorical factor analysis applied to ascertain common interpretable factors. The results have been presented to staff for discussion in the context of planning for improvement.

This paper explores the results of the factor analysis and its potential for providing academics with useful information on students’ experiences.

Effective teaching is more likely to be achieved by helping teachers to understand how to interpret research findings within their own context and circumstances, and so to identify the strongest influences on their own students. They will then be better able to think seriously about how their own practice can be enhanced in the light of the best research evidence currently available. (Entwistle, 2005, p. 81).

Theoretical background: the significance of disciplinary differences

In 1993, the Course Experience Questionnaire was introduced in Australia as a graduate survey, designed to collect data on the quality of students’ learning experiences in Australian universities. Since that time, while CEQ results have been used to measure university performance, annual data have consistently shown that they vary not by university but by discipline. There has been an overall improvement in the proportion of positive responses to CEQ items, but differences between disciplines have persisted.

The CEQ Good Teaching items focus primarily on the effectiveness of communication between academic staff and students. They were developed from work by Entwistle and Ramsden which demonstrated that effective communication with and by lecturers was strongly associated with students adopting a deep approach to their study and hence learning...
effectively (Ramsden, 1991, pp. 132, 135; Wilson et al, 1997, p. 43). These items were envisaged as a proxy measure of the quality of student learning.

The pilot data collected prior to the implementation of the CEQ as a graduate survey indicated strong and significant differences in the responses of students from different disciplines (Ramsden, 1991, p. 138). The persistence of differences between disciplines can readily be seen by inspecting the annual CEQ data, now available on the Graduate Careers Australia website (GCA 2008); see also Patrick (2005) and the graph at attachment 1. In general, social science and humanities disciplines are highly rated on the Good Teaching scale, whereas science and engineering disciplines are relatively poorly rated. Ramsden recognised this in his initial paper on the CEQ; he concluded that the differences between disciplines in terms of culture and resources are so marked that comparisons between institutions should only be made within disciplines (Ramsden, 1991, p. 139).

Differences between disciplines in terms of approach to teaching have emerged from other studies. Using different items, Santhanam and Hicks (2002) analysed the differences in students’ opinions in evaluating their lecturers and their subjects, across two principal discipline areas. The items in their study asked students to rate the effectiveness of the teacher and the effectiveness of the curriculum. In this study, sciences/mathematics students were more positive about the teaching they received than arts/humanities/social sciences students. Santhanam and Hicks concluded that content differences between disciplines influenced both the teachers’ approaches and the expectations of the students. This interpretation is consistent with Becher’s view of disciplinary culture; he argues that there are significant differences between disciplines in teaching techniques, student learning needs, and curriculum design (Becher, 1994). Similarly, Lueddeke argues that disciplinary differences have a significant influence on academics’ interest in scholarly reflection and teaching improvement (Lueddeke, 2004). Exploring teaching approaches, Trigwell, Prosser and their colleagues have also found disciplinary differences: they argue that expository teaching is more common in the “hard” sciences and a student-centred approach more common in social sciences and the arts (see eg Trigwell 1995; Lindblom-Ylänne et al., 2006).

Entwistle (2005) reports recent discipline-based investigations in the U.K. which complicate these assertions. He describes university education as a process of initiation into the complex culture of a particular field. While this does involve significant conceptual development (for which a “deep approach” is particularly useful), he now sees it more broadly as the development of a distinctive, discipline-specific way of “thinking and practising” (Entwistle, 2005, p. 72).

While recognising the significance of disciplinary cultures, Trowler and Knight (2000) also challenge the widespread practice of focusing on the practice of the individual teacher. They argue that attempts to improve students’ learning experiences must take account of contextual and structural issues, and the changes that they have observed in university life over the past twenty years. In particular, they nominate intensification of work; managerialism and a loss of academic autonomy; a loss of collegiality; “greedy” institutions; and ageing, malaise and marginality among academic staff (Trowler and Knight, 2000, pp. 71-72).

The present project is designed to explore the dimensions of student experience at discipline level. The aim is to provide academic staff with data which connect with the experiences of their own students and which will be useful to them in identifying where and how students’
experiences might be improved. This paper focuses in particular on strategies for presenting the analyses generated by the project.

**Methodology**

**Sample and survey**

It is university policy that each subject [locally termed a course] conducts a survey at least once a year. The data analysed in the study derive from the 2007 subject surveys; the data analysed for this paper were collected from surveys conducted in semester two (July - October). In this period, surveys were completed in 1466 Higher Education courses across the university, covering 50 discipline areas – a total of 33156 completed survey forms. Many students will have completed more than one survey form: fulltime enrolment is four subjects per semester, and fulltime students could well be asked to complete four surveys in a semester.

The survey instrument being used (the Course Experience Survey) has been locally developed to explore different aspects of student experience. While its statistical properties have not been previously evaluated, it does include the items of the CEQ Good Teaching scale and other items from the CEQ. It also includes items relating to study resources and learning facilities.

Table 1 lists the items in the survey, along with the labels used by the researchers. The Good Teaching items from the CEQ are highlighted. Responses (as with the CEQ) are on a 5-point Likert scale, with 1 labelled ‘strongly disagree’ and 5 labelled ‘strongly agree’.

**Table 1: Course Experience Survey items**

<table>
<thead>
<tr>
<th>Course Experience Survey item</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 The learning objectives in this course are clear to me</td>
<td>Objectives</td>
</tr>
<tr>
<td>2 I am learning what I expected to in this course</td>
<td>Expectations</td>
</tr>
<tr>
<td>3 This course is well organised</td>
<td>Organised</td>
</tr>
<tr>
<td>4 The teaching staff are extremely good at explaining things</td>
<td>Explaining</td>
</tr>
<tr>
<td>5 The teaching staff normally give me helpful feedback on how I am going in this course</td>
<td>Helpful feedback</td>
</tr>
<tr>
<td>6 This course contributes to my confidence in tackling unfamiliar problems</td>
<td>Problem-solving</td>
</tr>
<tr>
<td>7 Assessment tasks in this course require me to demonstrate what I am learning</td>
<td>Assessment</td>
</tr>
<tr>
<td>8 The amount of work required in this course is about right</td>
<td>Workload</td>
</tr>
<tr>
<td>9 The teaching staff in this course motivate me to do my best work</td>
<td>Motivate</td>
</tr>
<tr>
<td>10 I enjoy doing the work for this course</td>
<td>Enjoyment</td>
</tr>
<tr>
<td>11 I find the learning resources for this course useful (eg. notes, handouts, readings, AV materials)</td>
<td>Learning resources</td>
</tr>
<tr>
<td>12 The web-based (online) materials for this course are effective in assist my learning</td>
<td>Online materials</td>
</tr>
<tr>
<td>13 There is effective use of other computer-based teaching materials in this course</td>
<td>Computer-materials</td>
</tr>
<tr>
<td>14 The facilities (such as classrooms, lecture theatres, studios, labs) are adequate for this course</td>
<td>Facilities</td>
</tr>
<tr>
<td>15 I feel I can actively participate in my classes</td>
<td>Participate</td>
</tr>
<tr>
<td>16 There is good balance between theory and practice</td>
<td>Theory/prac balance</td>
</tr>
<tr>
<td>17 The teaching staff work hard to make this course interesting</td>
<td>Interest</td>
</tr>
</tbody>
</table>
### Data analysis

In previous studies, the robustness of the CEQ scales has been tested by factor analysis across the whole body of respondents. Implicitly, this approach assumes that the dimensions of students’ experience will be the same across different disciplines, and that what varies is the quality of their experience. By contrast, the present study uses factor analysis at discipline level, to distinguish the dimensions of student experience for each discipline (see Bedford et al., 2008).

Because responses to the survey items were on a Likert scale, which cannot be assumed to be linear (Meulman, Van der Kooij, & Heiser, 2004; Linting, Meulman, Groenen, & Van der Kooij, 2007), the data were first transformed using categorical principal component analysis (CATPCA). CATPCA is a non-trivial function of SPSS which commences analysis via optimal quantification, a process whereby categories of the categorical variables are appointed continuous numeric values (Linting et al., 2007; Meulman et al., 2004). This process provides the numeric values which are required for variance and Pearson correlation calculations (Linting et al., 2007). Importantly, the solution is iteratively computed from the ordinal data, as opposed to being derived from a correlation matrix, as with traditional PCA (Gifi, 1990). Like traditional PCA, CATPCA attempts to extract factors that can account for as much variance in the variables as possible. Because of the transformation of the fixed values into quantified values, CATPCA typically accounts for more variance than PCA (Linting et al., 2007).

Factor analyses were undertaken separately for each of the discipline areas for which data were available. The extraction method used was Principal Component Analysis, with Varimax Rotation and Kaiser Normalisation. Five factors were elicited, corresponding to the number of factors identified in traditional principal component analysis of CEQ data (see Ramsden, 1991; Richardson, 1994).

### Results

The five factor model had considerable explanatory power, accounting for around 70% of the variation in responses. An initial factor analysis across the whole cohort of students resulted in the emergence of a clear Good Teaching factor comprising five of the six expected items. At this level, the factors which emerged were:

- **Good teaching:** comment (Q20), give useful feedback (5), understand difficulties (19), motivate (9), interest (17)
- **Engagement:** use (18), enjoy (10), theory/prac balance (16), workload (8), problem-solving (6), participate (15), assessment (7)
- **Organisation:** objectives clear (1), organised (3), expectations (2), explaining (4)
- **Resources:** online (12), computer (13), learning resources (11)
- **Facilities:** facilities (14)

Analysis of variance using the established Good Teaching scale showed, as expected, strong and significant differences between disciplines, with mean scores ranging from 16 to 69 on a
scale from -100 to +100. As in previous studies, engineering and business students rated their experience low on this scale by comparison with social science and education students. The results are shown for some example disciplines in Figure 1, below.

**Figure 1  Student Good Teaching error bars, selected disciplines, Semester 2 2007**

![Graph showing student good teaching error bars for selected disciplines in Semester 2 2007.](image)

The factor analyses by discipline complicate and add depth to this picture.

While the items from the CEQ Good Teaching scale were commonly associated in a single factor, the significance of this factor varied between disciplines, and items associated with the factor also varied. In 26 of the 50 disciplines, a Good Teaching factor emerged first from the analysis; however it was the second factor in 16 disciplines and less salient in the remaining 8.

Table 2 provides an example comparison of the two most salient factors in two disciplines, one from the social sciences, the other from commerce. Mean scores on the standard GTS scale were relatively high for both these disciplines, as shown in Figure 1.

**Table 2 Example factor structures**

<table>
<thead>
<tr>
<th>Discipline</th>
<th>Communications N= 2573</th>
<th>Marketing N= 1451</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Factor 1</strong></td>
<td><strong>Good teaching</strong></td>
<td><strong>Organised and purposeful</strong></td>
</tr>
<tr>
<td></td>
<td>Questions 20 5 19 9 17:</td>
<td>Questions 3 4 2 1 17:</td>
</tr>
<tr>
<td></td>
<td>20 Staff put a lot of time into commenting on my work</td>
<td>3 This course is well organised</td>
</tr>
<tr>
<td></td>
<td>5 Teaching staff normally give me helpful feedback on how I’m going</td>
<td>4 The teaching staff are extremely good at explaining things</td>
</tr>
<tr>
<td></td>
<td>19 Staff make a real effort to understand any difficulties I might be having with my work</td>
<td>2 I am learning what I expected to in this course</td>
</tr>
<tr>
<td></td>
<td>9 Teaching staff motivate me to do my best work</td>
<td>1 The learning objectives in this course are clear to me</td>
</tr>
<tr>
<td></td>
<td>17 The teaching staff work hard to make this course interesting</td>
<td>17 The teaching staff work hard to make this course interesting</td>
</tr>
<tr>
<td></td>
<td>Factor accounts for 19% of variance</td>
<td>Factor accounts for 19% of variance</td>
</tr>
</tbody>
</table>
Despite the similarity in their GTS means, a good teaching factor emerged as the first factor for one of these disciplines but not the other. In communications, there was most variation in students’ experience of teaching, whereas in marketing, there was most variation in students’ experience of their subjects as purposeful and well organised.

Across the disciplines surveyed, differences in the items contributing to each dimension in different disciplines suggests slightly different perspectives on the meaning of the items. For example, responses to items about the organisation of the course (questions 1, 2, 3 and 4) were frequently correlated. In some disciplines these items also correlated with responses to Q18 (on the usefulness of the course), suggesting that the dimension related to whether the course was seen as purposeful. In other disciplines, they correlated with items on the requirements and outcomes of the course (questions 7, 8 and 6), suggesting that the dimension related to the course being seen as fit for purpose.

In addition to the diversity of the dimensions identified by the factor analysis, there was diversity in the salience of particular items within a factor. For example, among the Good Teaching items, Q20 (staff time on comment) was most frequently the defining item. However there were several disciplines where Q19 (understanding difficulties) or Q5 (useful feedback) was the defining item on this dimension.

**Reporting results**

Once the factors had been identified, the items contributing to each factor were grouped together and reports prepared for each discipline showing the items most closely associated with the “good teaching” factor. A sample image from the report is shown at Figure 2.

**Figure 2  Items associated with the Good Teaching dimension**
The item closest to the bullseye is the one loading most heavily on the Good Teaching factor. Additionally, each report includes a graph showing the level of agreement with each item and a table with the factor structure for the discipline.

The factor structure is also being used as the starting point for further work comparing student responses across subjects within the discipline: the project team has prepared graphs using error bars to show how students respond to associated items in different subjects.

However, communicating the results has been problematic. For staff unfamiliar with factor analysis, it is not obvious how to interpret an item which has the highest weighting on a factor, or how to interpret differences between factors.

The complexity of the factor analysis approach makes it particularly important to present the results using different perspectives. How might the key results of this work be presented so that the underlying logic is clear to a non-mathematician?

Three concepts seem to be fundamental.

1. The highest loading factor is the one which captures the greatest variability.
   This implies that students give a wide range of responses to the items in the factor. If their responses were alike there would be no variability to capture.

Factors are calculated to have a mean of zero and a standard deviation of 1, so factor scores per se cannot be used to demonstrate the variability of the underlying data. However, one would expect subject-level factor scores to vary and therefore to provide data illustrating the diversity within the discipline. Figure 3 depicts means and 95% error bars for subjects in the economics discipline (N=1207) (identifying detail removed).

   Figure 3  Error bars for Good Teaching factor scores, economics subjects

   As Figure 1 showed, Economics had a relatively low Good Teaching scale mean, of just over 20. Figure 3 shows that (as expected) the courses within Economics are not all alike. As

\[
\text{Figure 3  Error bars for Good Teaching factor scores, economics subjects}
\]
indicated by the 95% error bars, student responses are consistent within subjects - with the obvious exception of subject L, which had responses from only six students.

The circles in the error bars show the mean factor score for each course. There are clearly significant differences between courses, with some being rated much more positively than others. A couple of subjects (D and N) have a mean factor score nearly half a standard deviation above the mean, whereas two other subjects (C and H) have a mean score nearly half a standard deviation below the mean. D and N may offer examples of good practice; in C and H there appears to be considerable room for improvement. Hence analysis of factor (and item) scores by subject provides a useful starting point for further discussion.

2. The highest loading item on a factor reflects the variation in the other items.
This implies that there is variation in the responses to the highest loading item. It may be thought that the highest loading item on a factor will be one where there is strongest disagreement – the item which most needs fixing. This is not so.

This can be seen by looking at the distribution of responses to question 20 (teachers’ time commenting on student work). This was the defining item for the Good Teaching factor in a range of disciplines, including all those shown earlier. Although students were, as expected, much more positive in the arts and social science disciplines, in all the disciplines there was a broad range of responses.

Figure 4 Distribution of responses to Q20: Example disciplines

![Figure 4](image)

The defining item points to potential for examining where students' experiences are positive and where they might be improved. In the knowledge that this item varies significantly within the discipline, it is clearly valuable to develop more fine grained analysis looking at the experiences of different cohorts (international students, students articulating from TAFE, commencing and returning students). This work is being undertaken as part of the project.

3. Students respond consistently to the different items in the factor
An individual’s response on one item in the factor is likely to be similar to their response on the other items. A student who disagrees with one item will be likely to disagree with the other items; a student who agrees with one item will be likely to agree with the other items. Hence strongly disagreeing with an item in the factor will be associated with a low factor score and strongly agreeing will be associated with a high factor score.

Data from the project illustrate this. The graph below again draws on data from the economics
factor analysis. It shows the distribution of Good Teaching factor scores according to the response the students gave to question 20 (time on comments).

Figure 5  Error bars for Good Teaching factor scores, economics discipline, by response to Q20

It can be seen that within the discipline, students' experiences on this item are consistent with their experience of the other items included in the factor. The factor points to associated practices which can be considered as a set.

Discussion

Using the discipline level factor analysis provides more information about student experience than is visible from scanning item level agreements and scale means.

In Marketing, for example, Factor 1 comprised questions which seem to relate to students’ experiences of order and purpose in the subjects being evaluated (see Table 2: extract below).

<table>
<thead>
<tr>
<th>Extract from Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marketing</td>
</tr>
<tr>
<td>N= 1451</td>
</tr>
<tr>
<td>Factor 1: Organised and purposeful</td>
</tr>
<tr>
<td>Questions 3 4 2 1 17:</td>
</tr>
<tr>
<td>3 This course is well organised</td>
</tr>
<tr>
<td>4 The teaching staff are extremely good at explaining things</td>
</tr>
<tr>
<td>2 I am learning what I expected to in this course</td>
</tr>
<tr>
<td>1 The learning objectives in this course are clear to me</td>
</tr>
<tr>
<td>17 The teaching staff work hard to make this course interesting</td>
</tr>
<tr>
<td>Factor accounts for 19% of variance</td>
</tr>
</tbody>
</table>

Following through on the theoretical discussion, we can say first that this is the dimension of student experience where responses from Marketing students show most variation. Hence we know that on this dimension Marketing subjects include both good practice and opportunities for improvement. Some students must have indicated that they were clear about what they are doing, that it matched their expectations, and that they understood the work; other students (or
the same students in a different subject) must have indicated that they felt muddled about both the purpose and the content of the subject. It would be useful for staff to explore what is going on here. Where are the difficulties students experience? Are there particular groups of students who find the marketing subjects confusing? or are there particular subjects (perhaps problem-based subjects) where students are generally confused? Conversely: are there groups of students who are very clear about what they are doing and why? or are there particular subjects which are generally agreed to be organised and purposeful? Answering these questions (via further data analysis or via direct discussion with students) will provide a starting point for improvement.

Secondly, the factor analysis tells us that from the students’ perspective the items within this factor were related. Perceptions of order and purpose constituted a consistent dimension of the students’ experience. This gives more to work with than the results for each item individually. We can conclude that once staff know where improvement is needed, it will be useful for them to make a concerted effort to improve and align communication with students about the intentions, outcomes and content of the subjects which are causing difficulty.

In conclusion
This project has considerable potential for enhancing the usefulness of subject survey data. The results can be used to:

- identify distinct dimensions of students’ experience
- explore differences in students’ experience within a discipline
- identify subjects within the discipline which are exemplars of effective practice on a particular dimension

The disciplinary starting point focuses attention on the student's experience, rather than the teacher's performance. The factor analysis enables the identification of areas where students' experiences vary, so that within the body of practice in the discipline there is clearly both a need for improvement and positive exemplars. With discussion grounded at discipline and subject level, it should be possible to explore contextual and resource issues impacting on student experience, along the lines suggested by Trowler and Knight (2000). Used in this way, these analyses have the potential to provide a real jumping off point for change.

References
GRADUATE CAREER COUNCIL OF AUSTRALIA (2001), Code of Practice for the public disclosure of data from the Graduate Careers Council of Australia’s Graduate Destination Survey, Course Experience Questionnaire and Postgraduate Research Experience Questionnaire. Accessed via


Attachment 1
Illustrative means for CEQ Good Teaching, 2004 - 2007

![Graph showing mean GTS for different disciplines from 2004 to 2007. The graph compares Visual Arts and Crafts, Marketing, Economics, Communications and Media Studies, and Civil Engineering.]