Paper presented at AARE 2002 Brisbane 1st to 5th December 2002

MEASURING QUALITY OF STUDENT EXPERIENCES AT A UNIVERSITY USING A RASCH MEASUREMENT MODEL

Russell F. Waugh

Edith Cowan University

Key words: quality, quality of experiences, ideal perceptions of quality, university,higher education, Rasch measurement, self-reports, students

Running head: QUALITY OF EXPERIENCES

12 June 2002

Address correspondence to Dr Russell F. Waugh at

Edith Cowan University, School of Education, Pearson Street, Churchlands,

Western Australia, 6018

Abstract

The USA Community College Student Experiences Questionnaire, involving eight aspects of quality, was revised and re-written for a university in Australia. The eight quality aspects are: My Course, The Library, My Lecturers, Student Acquaintances, The Arts, The Sciences, Writing, and Vocations. Stem-items were created for each aspect and conceptually ordered-by-difficulty. Each of the stem-items was answered from two perspectives, Ideally, this is what I think should happen, and These are my experiences during this semester. The two response categories were Not at all, and Yes, on one or more occasions. The convenience sample was 363 students studying education at an Australian university and data were analysed with a Rasch measurement program (Simple Logistic Model). The difficulties of the items were calibrated and ordered from 'easy' to 'hard' on the same scale as the student measures of Quality. The final item sample was 52 (26 stem-items times 2). The proportion of observed student variance considered true was 0.81 and the proportion of item variance considered true was 0.92. The results supported most of the model behind the construct of Quality for the eight aspects in which the stem-items were ordered-by-difficulty and the difficulties of the items in the ideal perspective were easier than their corresponding difficulties in the experience perspective.

MEASURING QUALITY OF STUDENT EXPERIENCES AT A UNIVERSITY

USING A RASCH MEASUREMENT MODEL

Australian universities, like those in other countries, have been subjected to pressure to be accountable, to provide good service, and to be efficient in their use of resources. While this pressure came directly from the Australian government, it was indirectly part of a wider pressure on business and the community that occurred in Australia in the 1980s (see Hattie, 1990; Waugh, 2002). In 1988, the then Hawke Labor government, through the Minister for Employment, Education and Training (John Dawkins MP), asked universities to provide evidence of their performance through various performance indicators (Dawkins, 1988) and stated that some funding would be tied to the achievement of the indicators. This caused some tension in Australian universities at the time (see Hattie, 1990) but, nevertheless, the government's push for accountability and efficiency continued. Quality audits of universities were implemented in 1993, 1994, and 1995, with some funding ties (see Committee for Quality Assurance in Higher Education, 1995).

Globalisation affected Australian universities through international competition for students. International students studying at Australian universities were paying full-fees set by the government and these students began to demand quality in terms of teaching and resources provided for them. "We are paying for this" was a common statement made by the students. In the UK, a Charter of Education required students to be treated as customers by the universities (Department of Education, 1993), and this put more pressure on Australian universities for quality services (Harman, 1994). Australian students continued the pressure for quality, good teaching in lectures and tutorials, and the provision of services during the 1990s, as they raised quality problems. In 1999, the then Minister for Education, Training and Youth Affairs (Dr David Kemp MP, a former academic) stated that an Australian Quality Assurance Agency would be set up in 2001 (Kemp, 1999). This Agency has now been set up and is planning some quality investigations from 2002 onwards.

All this continues the pressure on universities to provide quality services (like good teaching, computer facilities, library resources, lecture and tutorial facilities, academic advisers, counselling, and medical facilities) and quality results (like high completion rates, low costs

per student, good research results, many top PhD students, and high student satisfaction). However, although the universities can provide good services, at least part of the quality output must come from the students. Students must be prepared to put in the effort and take advantage of the services offered in order for the quality to be high. Quality is not just what the universities do, but also what the students do.

This leads to the problem for this paper, the measurement of quality of student experiences at a university. The purpose of the present study was to test a model of a construct of Quality of Student Experiences at a university. Stem-items were designed specially for each of eight aspects of Quality, in an ordered pattern from 'easy' to 'hard', based on the Community College Student Experiences Questionnaire (Friedlander, Pace & Lehman, 1990). For each stem-item, it was expected that the two response perspectives would also be ordered from 'easy' (Ideally, this is what I think should happen) to 'harder' (My experiences during this semester) for each stem-item. It is 'easier' for students to have high ideals of quality experiences at university than it is for students to put in the effort to take part in all the experiences that a university has to offer. Data were analysed with the computer program Rasch Unidimensional Measurement Models (RUMM) (Andrich, Sheridan, Lyne & Luo, 2000) using the Simple Logistic Model of Rasch.

For Quality of Student Experiences at university, there is a need to measure Ideal Perspectives and Perspectives by Experience on the same scale, instead of measuring them separately and then calculating the correlations between them. The latter method uses separate scales with different items for each perspective, that are usually based on Classical Test Theory, that are not at the interval-level and that are often not sufficiently reliable. There is a need to reduce the 'noise or error' in measures by producing a proper linear scale and by calibrating the difficulties of all the items on the same scale as the measures.

Previous measurements of quality of student experiences

Pascarella and Terenzini (1991) reviewed the literature on the effects of college on USA students. They made three main conclusions: (1) there are many within-college experiences that maximize influence on students and many of them are independent of the college attended; (2) most of the experiences that maximize influence on students are dependent on the students; and (3) 'the impact of college is largely determined by the individual's quality of effort and level of involvement in both academic and non-academic activities' (p.610). This conclusion implies that the impact of college is not totally dependent on what the college does for, or to, students, although this is important too. A substantial part of the impact (according to Pascarella and Terenzini, 1991, pp.610-611) depends on the extent to which students involve themselves with other students, with lecturers (professors), and with the programs, research talks, laboratory sessions and extra-curricula activities (such as debates, political discussions, plays, music recitals, science issues and student groups) that are available at the college (or university).

Friedlander, Pace and Lehman (1990) developed the Community College Student Experiences Questionnaire to measure the quality of effort and experiences of students at USA colleges, based on the work of Pace (1979a, 1979b, 1984) (for which norms were published by Pace (1992), and a test manual by Lehman (1991, 1992)). This questionnaire has eight sub-scales relating to student effort and experiences in: (1) My Course, (2) Library usage, (3) Faculty, (4) Student acquaintances, (5) The Arts, music and theater, (6) Writing, (7) The Sciences, and (8) Vocations. Using Classical Test Theory, maximum-likelihood factor analysis with oblique rotation, and several large samples of several thousand students from across the USA, Ethington and Polizzi (1996) concluded that data collected with the Community College Student Experiences Questionnaire had good construct validity and eight factors, as postulated, and that the eight factors 'explained' 53% of common variance

amongst the items. It was suggested that the questionnaire could be 'used to make valid and reliable inferences regarding students' efforts and involvement' at colleges across the USA (Ethington & Polizzi, 1996, p.711).

Koljatic and Kuh (2001), in response to government and industry requests to improve the quality of college and university education in the USA, investigated data from 73, 050 students who had completed the College Student Experiences Questionnaire between 1983 and 1997. They concluded that there was minimal variation, with random fluctuations, in the quality of student engagement in three good educational practices (cooperation with peers, active learning and faculty-student interaction) in this 15-year period.

In Australia, universities use the Course Experience Questionnaire to measure student satisfaction with their degree course and the teaching of it, four months after graduation (Ainley, 2001a; Waugh, 1998, 1999; Wilson, Lizzio & Ramsden, 1997). The reason for this timing is that it is administered with a Graduate Destination Survey (involving type of job, employer and salary details). The Course Experience Questionnaire has five sub-scales (25 items with five Likert response categories SDA, DA, N, A, SA): Good teaching, Clear goals and standards, Appropriate assessment, Appropriate workload, and Generic skills, and a single item on Overall satisfaction with the course. For postgraduates, Australian universities use the Postgraduate Research Experience Questionnaire (Ainley, 2001b). This contains six sub-scales (28 items with five Likert response categories): Supervision (supervisors), Thesis examination process, Skills development, Goals and standards, Intellectual climate, Infrastructure (working space and technical support), and a single item on Overall satisfaction. Both these questionnaires only measure a part of the experiences of students at university and they focus more on satisfaction rather than on the quality and extent of student experiences in the wider aspects of a university.

Waugh (2001) reported on a study in which the USA Community College Student Experience Questionnaire was revised and used with a sample of 305 first-year undergraduates at one Australian university to measure student quality, effort and experiences. Four main revisions were made: (1) All items were re-written in a positive format under the eight aspects of quality; (2) All 60 items were answered in two perspectives (expectations at the beginning, and experiences during, the course); (3) The items were ordered under their respective eight aspects so that it was clear what was being measured; and (4) The response format was replaced by one involving applicability to the number of subjects studied. The data were analysed with a Rasch measurement model computer program. It was found that 58 of the 120 items didn't fit the measurement model and that most of these came from The Arts, The Sciences and Writing. It was concluded that data from the other 62 items had good reliability and construct validity, and that expectations were easier than experiences for all items.

Problems with the Community College Student Experiences Questionnaire

As used in the USA by Friedlander, Pace and Lehman (1990), Ethington and Polizzi (1996), and Koljatic and Kuh (2001), there are at least four aspects of the Community College Student Questionnaire that are called into question. One, researchers may not be aware that Quality of Experience can be thought of as a model based on a uni-dimensional trait involving a philosophy of selecting data to fit a Rasch measurement model. One can have a scale of Quality of Experience with measures from low to high, calibrated on the same scale as the item difficulties from 'easy' to 'hard'. Two, no check is made that the conceptually ordered response categories are answered consistently and that students can discriminate between the categories consistently. Three, confirmatory and exploratory factor analysis, and correlation methods do not calibrate the item difficulties on the same scale as the Quality of Experience measures and items that are measuring 'noise' or 'something else' can

be inadvertently left in the measure to distort the correlations. Four, the methods used imply that the variables related to Quality of Experience (like ideal perceptions and actual experiences) cannot be measured on the same scale. It should be possible, for example, to measure ideal perceptions and actual experiences on the same scale for all eight aspects.

What is needed is a way of creating a Quality of Experience scale in which 'noise' is reduced considerably and the difficulties of the items of Quality are calibrated together on the same scale with the student measures of Quality of Experiences. This can be done with a Rasch Measurement Model (see Andrich, 1988a, 1988b, 1985, 1982; Rasch, 1960/1980) and it requires a different kind of questionnaire from the usual Likert format for self-report data.

Aims

The aims of this study are to:

1. To test a structure of Quality of Student Experiences based on eight aspects (My Course, The Library, My Lecturers, Student Acquaintances, The Arts, The Sciences, Writing and Vocations), and two perspectives (Ideally, this is what I think should happen and My experiences during the semester);
2. Create a questionnaire to measure Quality at university based on the structure given above in which stem-items for each aspect are ordered-by-difficulty;
3. Create a new scale of Quality at university in which the item difficulties are calibrated on the same scale as the student measures of Quality, using the Rasch Unidimensional Measurement Model (RUMM) computer program (Andrich, Sheridan, Lyne & Luo, 2000).

Model of the Construct: Quality of Student Experiences

Initially, sets of stem-items were created for the eight aspects of Quality of Student Experiences. All the stem-items were created and written especially for the present study, based on items and ideas from Friedlander, Pace and Lehman (1990) and Waugh (2001). They were written with the intention that they would be seen to be ordered in increasing difficulty. An example is now given (see also Appendix A).

Stem-Items for My Course: Discussions (Ordered by conceptual difficulty)

It was conceptualized that most students would probably find it very easy to say that, ideally, they think that they should ask questions about points made in class discussions, lectures and readings (stem-item 4) during the semester. It was conceptualized that most students would probably find it harder (but still very easy) to say that, ideally, they should participate in class discussions and discuss different points of view (stem-item 5), because it involves a 'little bit more' than stem-item 4. That is, stem-item 5 involves the content of stem-item 4 plus the process of discussing different points of view that makes it harder than stem-item 4. It was conceptualized that most students would probably find it harder to say that ideally they should discuss and explain new material, and argue the force of different points of view (stem-item 6), because it involves a 'little bit more' than stem-item 5. That is, stem-item 6 involves the content of stem-item 5 plus the process of arguing the force of different points of view that makes it harder than stem-item 5. It was conceptualized that most students would probably find it harder to say that, ideally, they should do additional readings on topics that were introduced in class or lectures and explain these to other students (stem-item 7), because it involves a 'little bit more' than stem-item 6. That is, stem-item 7 involves the content of stem-item 6 plus the process of doing additional work and explaining it to others that makes it harder than stem-item 6. So it was expected that these four stem-items would form an ordered-by-difficulty pattern from easy to hard, when students reported on them

as ideally, this is what they think should happen, and there might be some variation around this.

It was conceptualized that this same ordered-by-difficulty pattern of responses for stem-items 4 (very easy), 5, 6, to 7 (hardest, but still easy) would apply when students reported on their experiences during this semester, except that they would be correspondingly harder than their views of ideally, this is what they think should happen. These patterns can be seen in the questionnaire (see Appendix A). That is, the difficulties of stem-items 4-7 are conceptually ordered vertically from easy to hard for both the ideal and actual perspectives and, for each stem-item, the difficulty of the two perspectives, ideal and actual, are ordered from easy to hard, horizontally.

Expected Ordered-by-Difficulty Patterns for the Other aspects

The conceptualized ordered-by-difficulty patterns for the other stem-items follow a similar line to that reported above and are not reported here to avoid repetition. A reader can easily work out the other patterns from Appendix A. The stem-items are placed in conceptual order from easy to hard in each aspect vertically and the two perspectives (Ideally, this is what I think should happen and My experiences during this semester) are placed in conceptual order from easy to hard, horizontally. Calculated item difficulties in logits are included.

Possible Weakness in Item Wording

There is a potential weakness in the wording of the items as they are designed contrary to the usual advice given in item construction. That is, items should be kept simple with one main idea so that respondents can clearly focus on the one idea, and both respondent and designer know what is being measured. When the wording of items is complicated by the inclusion of more than one main idea, it could be argued that different respondents focus on different ideas in the same item, and so measurement is compromised. Since some of the items in the present scale contain two or three verbs or nouns, and more than one main idea, this argument leads to the view that good measurement is unlikely with the present scale. The counter view is that some items are constructed with more than one idea in order to make them conceptually harder in an ordered-by-difficulty pattern and thus build-in the idea of measurement from easy to hard for the items (and from low to high for respondent abilities) into the scale. The judge of whether a good scale is produced is whether the data support the conceptual design.

Method

Administration and sample

The questionnaire was pre-tested with six students and discussed with them before being used in the study. Some changes were made to the wording, as considered appropriate. The student-sample consisted of 363 students at an Australian University studying in Education and is a convenience sample. There were 24 (6.4%) in 4thyear BEd at Campus 1, 72 (19.25%) in 4th year BEd at Campus 2, 152 (40.6%) in 4th year BEd at Campus 1, and 126 (33.7%) in 4th year BEd at Campus 2.

Permission was obtained from the University Ethics Committee and from the lecturers of university classes to administer the questionnaire during classes. The questionnaire structure is more complex than students are used to answering and this had to be explained to them. The author remained with the students after the explanation, in order to explain problems of interpretation as they arose. Nine questionnaires had to be discarded because

students answered wrongly by omitting many questions or not understanding the structure. Students generally took 15-20 minutes to complete the questionnaire.

Measurement Model

The Simple Logistic Model of Rasch (Andrich, 1988; Rasch, 1980/1960) was used with the computer program Rasch Unidimensional Measurement Models (RUMM) (Andrich, Sheridan, Lyne & Luo, 2000) to analyze the data. Items fitting the model were calibrated from easy to hard and student measures were aligned from low to high on the same scale. The Rasch method produces scale-free student measures and sample-free item difficulties (Andrich, 1988; Wright & Masters, 1982, 1981). This means that the differences between pairs of student measures and pairs of item difficulties are expected to be sample independent - one of the requirements of measurement.

The Rasch model requires that data must fit the measurement model and not the other way round (see Andrich, 1989). This follows from the requirements needed to create a proper scale. There are four main requirements (see Wright, 1999, p.100). One is that the measures should be linear so that inferences of additivity, separability and divisibility can be used in practice. A second is that item calibrations must be sample-free. That is, a proper scale is invariant across groups for which it is used. This means that, for the Rasch model, all the items contributing to the scale must have the same discrimination parameter. A third is that the created scale should not be affected by the opinions of students whose answers are used to construct it. A fourth is that it should be possible to omit some items without affecting an individual's measure on the scale.

The RUMM program (2000) calculates standard errors of measurement for the student measures of Quality and for the item difficulties, as well as a Student Separability Index. The equations for these are given in Wright and Masters (1982). The Index shows the proportion of observed variance considered true.

The RUMM program substitutes the parameter estimates back into the model and examines the difference between the expected values predicted from the model and the observed values. The program provides individual item and student fit statistics and two global tests-of-fit. One is the item-trait interaction and the second is the item-student interaction. The item-trait test-of-fit (a chi-square) examines the consistency of the item parameters across the student measures for each item and data are combined across all items to give an overall test-of-fit (see Andrich & van Schoubroeck, 1989, pp.479-480 for the equations). This shows the collective agreement for all items across students of different Coping measures. The item-student test-of-fit examines both the response patterns for students across items and for items across students. It examines the residual between the expected estimate and the actual values for each student-item summed over all items for each student and summed over all students for each item (see Styles & Andrich, 1993, p.914 or Andrich & van Schoubroeck, 1989, p.482 for the equations). The fit statistics approximate a distribution with a mean near zero and a standard deviation near one, when the data fit the measurement model. Negative values indicate a response pattern that fits the model too closely (probably because response dependencies are present, see Andrich, 1985) and positive values indicate a poor fit to the model (probably because other measures -'noise'- are present).

Results

The results are set out in Appendices A and B, Tables I, II and III, and Figure 1. Appendix A shows the items and the item difficulties. Appendix B shows the difficulty, the standard error, the residual and the fit to the measurement model for each item. Table I provides a summary of the global statistics for the Quality of Student Experiences scale data. Table II shows the

agreement in rank between the expected and actual item difficulties within each sub-group (vertical ranking). Table III shows the agreement in rank between the expected and actual item difficulties within each sub-group (horizontal ranking). Figure 1 shows a graph of the item difficulties on the same scale as the student measures of Quality of Student Experiences.

Place Figure 1 about here

Characteristics of the Data

Psychometrics

The Index of Student Separability for the 52-item scale with two response categories is 0.81. This means that the proportion of observed student variance considered true is 81% and that the errors in Quality measures (about 0.38 logits on average) are small in relation to the separation of the Quality measures along the scale (SE= 1.26 logits). The power of the tests-of-fit, based on the Student Separability Index of 0.81, is good. The Index of Item Separability is 0.92. This means that the proportion of observed item variance considered true is 92% and that the errors in item difficulties are small (about 0.28 logits on average) in relation to the separation of item difficulties along the scale (SE=1.79 logits).

The 52 items of the Quality Scale have a reasonable global fit to the measurement model (see Table I). The item-trait tests-of-fit indicate that the values of the item difficulties along the scale are consistent across the range of student Quality scores (Chi-square = 248, df=250, p=0.52) (see Table I). The null hypothesis is that there is no interaction between the responses to the items and the locations of the students along the trait. This means that students with different measures of Quality agree with each other as to which items are 'easy' and which are 'hard', as required to make a good scale.

The standardized mean-square residual has a mean near zero (-0.50 for items and +0.37) for students, and a SD near 1 (1.00 for items and 0.69 for students). These item-student tests-of-fit (see Table I) indicate that there is good consistency of student and item response patterns and that there is a good fit to the measurement model.

The difficulties of the items are not as well targeted against the students as they could be. The Quality measures range from about zero to +6.4 logits. The item difficulties range from about -4 to +3.8 logits (see Figure 1). So there are 57 students not 'covered' by the item difficulties in the upper part of the scale. There is a need to add some hard items that have difficulties in the region +4 to +7 logits and there are some stem-items that could be deleted because they are too easy for the students surveyed (such as 1, 2, 3, 4, 5 , 6 and 7).

The individual item fits to the measurement model are reasonably good (see Appendix B). The item residuals are all between than +/-2, except for items 26 (2.34) and 58 (-2.03). The item chi-squares for the item-trait interaction are all less than 11 (except for item 46) and their corresponding probabilities of fit to the measurement model are all greater than 0.04, except for item 46.

Place Table I about here

Agreement with the conceptual design

Agreement between the actual vertical ordering of the difficulties of the stem-items and the conceptual ordering of the difficulties is reasonably good (see Table II). In the Experience Perspective, only the difficulties of stem-items 4 and 5 (under My Course: Discussions) were disordered. It is difficult to see why this is so. In the Ideal Perspective, the difficulties of stem-items8 and 9, 11 and 13, 16 and 17, and 28 and 29, are equal within their standard errors. This need not mean that these stem-items should be reworded. They were ordered in the Experience Perspective and their data fitted the measurement model. Data from stem-items 12, 19, 21 and 24 did not fit the measurement model and were discarded. The main reason for non-fit was that students couldn't agree on their difficulties. Many students with high quality measures answered no on these stem-items and many students with low quality measures answered yes. These stem-items do not need to be replaced or reworded in any future use of the scale, but they could be.

Place Table II about here

Agreement between the actual horizontal ordering of the difficulties of the stem-items and their conceptual ordering by perspective is as expected (see Table 4). That is, students reported that their Ideal Perspectives were easier than their Perspectives by Experience by more than 2 SEs for all stem-items.

Place Table III about here

Discussion

The model of Quality of Student Experiences

The data analysis supports the view that eight aspects of Quality of Student Experiences at a university are linked together. These are: (1) My course (2) The library, (3) My lecturers, (4) Student acquaintances, (5) The Arts, (6) The Sciences, (7) Writing, and (8) Vocations. Stem-items from all eight aspects were placed by difficulty on the same scale and shown to be associated with one dominant trait, the measure of Quality of Experiences. There is strong, but not total, support for the conceptualized ordered structure behind the model of Quality of Student Experiences. There is strong support for the view that students' Ideal Perspectives are easier than their Perspectives by Experience for items in each of the eight

aspects. All the evidence provided supports the conclusion that the data have good construct validity and are internally consistent and reliable.

Interpretation

The 52 items of the scale are ordered from 'easy' to 'hard' and the item difficulties form a scale from which inferences of linear measurement of Quality of Student Experiences can be made (see Figure 1). Nearly all the students answered the 'easy' items positively (for example, see the ideal items 1, 2, 7, 9, 5, 19 in Table 1). As the item 'difficulty' becomes progressively higher on the scale, the students need a corresponding higher measure to answer them positively. The 'hardest' items are only likely to be answered positively by students who have high measures (for example, the experience items 34, 46, 32, 26, 40 in Table 1). Students with low measures of Quality are unlikely to be able to answer these 'difficult' items positively.

The three 'easiest' items in the Ideal Perspective are, in order: Combine ideas from different sources of information in preparing assignments (item 1, 'very easy'),Combine ideas, summarise major points and seek new information in preparing assignments (item 2), and Ask questions about points made in class discussions, lectures or readings, (item 7, 'hardest', but still 'very easy'). These three items were 'very easy' in their Perspective by Experience mode too, and they were 'easier' in the Ideal Perspective than in the Experience Perspective, as expected.

The three 'hardest' items in the Ideal Perspective mode are, in order: Attend something on The Arts at university (an art exhibition, concert, play, or theatre production) and discuss it with students (item 43,'hardest and reasonably hard'), Have serious discussions with students about the Sciences (eg pollution, genetics, technology or cosmology) (item 45), and Discuss my career plans, education plans, interests and ambitions with my lecturer (item 33, 'easiest', but still ' hard'). These three items are 'very hard' in their Perspective by Experience mode, and 'harder' than in the Ideal mode, as expected.

Implications

The results supported research by Friedlander, Pace and Lehman (1990) and Ethington and Polizzi (1996) on the Community College Student Experiences Questionnaire in the USA in that quality is associated with the same eight aspects of university life. However, the present study shows that many of the USA items are not part of a uni-dimensional measure. An implication is that the USA data analyses were not sufficiently stringent to create a proper linear measure from which measurement inferences can be made with confidence.

The findings about the impact of college being largely determined by the students' quality of effort and level of involvement in academic and non-academic activities (as suggested by Pascarella and Terenzini, 1991) are likely to be true about students in Australian universities. This is because the same eight aspects of university life in the USA quality scale were found to be associated with one dominant quality trait with Australian data (even though half the USA items had to be discarded).

An implication for researchers is that there should be more educational studies in which linear measures are made of important variables. Wright (1999) claims that one cannot produce linear measures in education and educational psychology and make proper inferences (1) by just using raw scores totalled from a number of items, (2) by allowing guessing parameters in the measurement model, (3) by allowing item discriminations in the measurement model to vary, (4) by not minimizing residuals, (5) by not creating measures that can be added and subtracted (see also Wright, 1985), and (6) by destroying construct

validity. Wright (1999) argues that the Rasch measurement model (as used in the current study to produce a linear scale) is producing a 'revolution' in educational measurement and its use can lead to the creation of 'stable and reproducible laws like we have in physics' (p.101). A further implication is that the measure of Quality created in the present study might be useful, along with other linear measures, in producing a 'law' in education. This would be an exciting development.

In the present study using a Rasch analysis, all the items were calibrated on the same scale together so that their 'difficulties' in relation to one another can be seen (see Appendix A and Figure 1) and so that the relationships between the items can be measured. The following example is a little complicated because the measure of Quality has no true zero. The difference between students actually discussing career plans, education plans, interests and ambitions with their lecturer and them feeling that ideally this is what should happen at university is about 7 times as hard as the difference between students actually asking their lecturers for information about grades, assignments and course work, and them feeling that ideally they should discuss their career plans, education plans, interests and ambitions with their lecturer [3.83 - (0.32)] / [0.81 - (0.32)].

In a second example, the difference between students actually asking a librarian for help in finding various books, journals and other material and them feeling that ideally this should happen is about twice as hard as the difference between students actually using library computers to locate up-to-date books and journals for assignments and study, and students ideally thinking that they should be asking a librarian for help in finding various books, journals and other material [2.39 - (-0.97)] / [0.64 - (-0.97)].

In a third example, the difference between students actually having serious discussions with other students about The Sciences and students thinking that ideally this should happen is about 1.2 times as hard as the difference between students actually attending something on The Arts at university and discussing it with other students, and students ideally thinking that they should have serious discussions with other students about The Sciences [3.53 - (0.77)] / [3.02 - (0.77)].

There are implications for university administrators and for students from the measurement of Quality of Student Experiences. The measurement suggests that there are many aspects students think should happen ideally, but that are difficult to actually experience. Some examples relate to The library, Student acquaintances, The Arts, The Sciences, writing and re-writing assignments, and practicing and explaining a vocational task. The university provides opportunities for these experiences but the extent to which students make use of them is up to the students and it is clear that many students do not make the most of their opportunities. Perhaps the university should inform students in the orientation week and in brochures of the benefits to be gained from taking part in the many experiences that are offered. It is highly probable that many students do not know or understand that the impact of university on them is largely determined by their quality of effort and level of involvement in both academic and non-academic activities (Pascarella & Terenzini, 1990; Ethington & Polizzi, 1996; Waugh, 2001). Some may 'know' this but still not appreciate the significance of it.

Students reported that discussing aspects such as career plans, education plans, interests, ambitions, current events and university issues with their lecturers were very hard to experience. This does not necessarily translate into implications for university administrators and lecturers. It is probably impossible for lecturers to provide enough time to discuss all these issues with all their students. It is more likely that lecturers would do this with their PhD students than with their undergraduate students. Nevertheless, an implication is that some students do experience these aspects and the implication is that they benefit from it.

The measure implies that there are some aspects where students are taking advantage of their opportunities and these relate to items concerning their course. Many of the items relating to combining ideas from different sources in preparing assignments, and asking questions and participating in classes and tutorial groups were very easy in the experience mode. This means that most students were able to answer them positively.

References

AINLEY, J. (2001a) **Course Experience Questionnaire 2000** (Parkville, Victoria: Graduate Careers Council of Australia).

AINLEY, J. (2001b) **The 1999 Postgraduate Research Experience Questionnaire** (Canberra: Department of Education, Training and Youth Affairs).

ANDRICH, D. (1989) Distinctions between assumptions and requirements in measurement in the social sciences. In J.A.KEATS, R.TAFT, R.A.HEATH, & S.LOVIBOND (Eds), pp. 7-16, **Mathematical and Theoretical Systems** (Elsevier Science Publishers: Amsterdam (North-Holland)).

ANDRICH, D. (1988) **Rasch models for measurement**. Sage university paper on quantitative applications in the social sciences, series number 07/068. (Newbury Park, California: Sage Publications).

ANDRICH, D. (1985) A latent trait model for items with response dependencies: Implications for test construction and analysis. In S.E. EMBRETSON (Ed), **Test design: developments in psychology and psychometrics** *(*pp. 245-275*)* (Orlando: Academic Press).

ANDRICH, D. (1982) Using latent trait measurement to analyse attitudinal data: a synthesis of viewpoints. In D. SPEARITT (Ed) **The improvement of measurement in education and psychology**, pp. 89-126. (Melbourne: Australian Council for Educational Research).

ANDRICH, D., SHERIDAN, B., LYNE, A. & LUO, G. (2000) **RUMM: A windows-based item analysis program employing Rasch unidimensional measurement models** (Perth: Murdoch University).

ANDRICH, D. & van SCHOUBROECK, L. (1989) The General Health Questionnaire: a psychometric analysis using latent trait theory, **Psychological Medicine**, 19, pp.469-485.

COMMITTEE FOR QUALITY ASSURANCE IN HIGHER EDUCATION (1995) **Report on 1994 quality reviews** (Canberra, AGPS).

DAWKINS, J.S. (1988) **Higher Education: A Policy Statement** (Canberra, AGPS).

DEPARTMENT of EDUCATION (1993) **The Charter for Higher Education** (London, HMSO)

ETHINGTON, C.A. & POLIZZI, T.B. (1996) An assessment of the construct validity of the CCSEQ Quality of Effort Scales, **Research in Higher Education***,* 37, 6, pp. 711-730.

FRIEDLANDER, J., PACE, C. & LEHMAN, P. (1990) **Community College**

Student Experiences Questionnaire (Memphis: University of Memphis, Centre for the Study of Higher Education).

HARMAN, G. (1994) Australian higher education administration and the quality assurance movement, **Journal of Tertiary Education Administration**, 16 (1), pp. 25-43.

HATTIE, J. (1990) Performance indicators in education, **Australian Journal of Education**, 34 (3), pp. 249-276.

KEMP, D. A. (1999) **Quality assured: A new Australian Quality Assurance Framework for university education**. Speech given by the Hon. Dr D. A. Kemp, Federal Minister for Education, Training and Youth Affairs, at the seminar on the New Quality Assurance Framework, in Canberra, Australia, 10 December 1999.

KOLJATIC, M. & KUH, G.D. (2001) A longitudinal assessment of college engagement in good practices in undergraduate education, **Higher Education,** 42 (3), pp. 351-371.

LEHMAN, P.W. (1992) **CCSEQ: Test manual and comparative data** (Los Angeles, CA: University of California, Centre for the Study of Evaluation).

LEHMAN, P.W. (1991) **Assessing the quality of community college student experiences: a new measurement instrument**. Unpublished doctoral dissertation. (Los Angeles, CA: University of California, Graduate School of Education).

PACE, C.R. (1979a). **The College Student Experiences Questionnaire** (Bloomington: Indiana University, School of Education).

PACE, C.R. (1979b) **Measuring outcomes of College: Fifty years of findings and recommendations for the future** (San Francisco: Jossey-Bass).

PACE, C.R. (1984) **Measuring the quality of student experiences** (Los Angeles, CA: University of California, Graduate School of Education, Higher Education Research Institute).

PACE, C.R. (1992) **College Student Experiences Questionnaire: Norms for the third edition, 1990** (Los Angeles, CA: University of California, Centre for the Study of Evaluation).

PASCARELLA, E.T. & TERENZINI, P.T. (1991) **How college affects students:Findings and insights from twenty years of research** (San Francisco, CA: Jossey-Bass).

RASCH, G. (1980/1960) **Probabilistic models for intelligence and attainment tests** (expanded edition) (Chicago: The University of Chicago Press) (original work published 1960).

WAUGH, R.F. (2002) Academic staff perceptions of administrative quality at universities, **Journal of Educational Administration**, 40 (2), pp. 172-188.

WAUGH, R.F. (2001) Quality of student experiences at university: A Rasch measurement model analysis, **Australian Journal of Education**, 4*5* (2), pp.183-206.

WAUGH, R. F. (1999) A revised Course Experience Questionnaire for student evaluation of university courses. In: **The Course Experience Questionnaire Symposium 1998** T.HAND

& K.TREMBATH (Eds), pp.61-80, (Canberra: Evaluations and Investigations Programme, Higher Education Division, DETYA).

WAUGH, R.F. (1998) The Course Experience Questionnaire: A Rasch measurement model analysis, **Higher Education Research and Development***,* 17 (1), pp. 45-64.

WILSON, L. LIZZIO, A. & RAMSDEN, P. (1997). The development, validation and application of the Course Experience Questionnaire, **Studies in Higher Education**, 22 (1), pp. 33-53.

WRIGHT, B.D. (1999) Fundamental measurement for psychology. In: S.E. EMBRETSON & S.L. HERSHBERGER (Eds) **The new rules of measurement: What every psychologist and educator should know**, pp. 65-104. (Mahwah, NJ: Lawrence Erlbaum Associates).

WRIGHT, B.D. (1985) Additivity in psychological measurement. In: E.E. ROSKAM(Ed), **Measurement and Personality Assessment**., pp. 101-112 (Amsterdam: North-Holland:Elsevier Science Publishers B.V.).

WRIGHT,B. & MASTERS, G. (1982) **Rating scale analysis: Rasch measurement** (Chicago: MESA Press).

WRIGHT, B. & MASTERS, G. (1981) **The measurement of knowledge and attitude** (Research memorandum no. 30) (Chicago: Statistical Laboratory, Department of Education, University of Chicago).

Appendix A

QUESTIONNAIRE: Quality of Student Experiences at University

This questionnaire is anonymous. Please don't put your name or any identification on it. Please read the consent and cover page.

Please rate the 60 statements according to the following perspective and place a number corresponding to **Ideally, this is what I think should happen** and **My experiences during the semester** on the appropriate line opposite each statement:

On one or more occasions put 2

Not at all (never) put 1

Example

If **Ideally, you do think that** you would never have to combine ideas from different sources of information in preparing assignments, put 1. If your **experiences during the semester** were that you had to combine ideas from different sources of information in preparing assignments **on many occasions**, put 2.

Item 1-2 Combine ideas from different sources 1 2

Item no. Item wording Ideally, this My experiences is what I think during this

should happen semester

Sub-Group: My Course (18 items)

Information and assignments

1-2 Combine ideas from different sources of

information in preparing assignments. <-4 -3.45

3-4 Combine ideas, summarise major points and

seek new information in preparing assignments. <-4 -1.71

6. Seek new information and consider the evidence

for it, in preparing assignments. -2.43 -0.56

Discussions

7-8 Ask questions about points made in class

discussions, lectures or readings. -3.63 -0.15

10. Participate in class discussions and discuss

different points of view. -2.96 -1.70

11-12 Discuss and explain new material, and argue

the force of different points of view. -1.72 +0.63

14. Do additional readings on topics that were

introduced in class or lectures and explain

these to other students. -0.57 +1.74

Practical

16. Have serious discussions with students

about practical aspects of my course. -1.88 +0.35

18. Attend a practical session relating to my course

(laboratory, practicum or demonstration). -1.89 +1.38

Sub-Group: The Library (8 items)

20. Use the library computers to locate up-to-date books and journals for assignments and study. -2.42 +0.64

22. Use the library as a quiet place to read and study. -1.08 +1.62

24. Use the library computers to browse and locate books and journals with interesting material. No fit No fit

26. Ask a librarian for help in finding various books, journals and other material. -0.97 +2.39

Sub-Group: My Lecturers (8 items)

28. Ask my lecturers for information about grades, assignments and course work. -1.81 +0.81

30. Talk (briefly) with my lecturers after class about course work and assignments. -1.33 +1.90

32. Discuss (briefly) current events, recent topics and university issues with my lecturer. +0.29 +2.83

33-34 Discuss my career plans, education plans, interests and ambitions with my lecturer. +0.32 +3.83

Item no. Item wording Ideally, this My experiences is what I think during the should happen semester

---

Sub-Group: Student Acquaintances (6 items)

36. Have serious discussions with students who are older and some who are younger than me. -1.08 +0.99

38. Have serious discussions with students whose ethnic, religious or cultural background is different from mine. No fit No fit

40. Have serious discussions with students whose political opinions or beliefs are different from mine. +0.17 +2.28

Sub-Group: The Arts (4 items)

42. Have serious discussions with students about The Arts (art, painting, music or theatre). No fit No fit

44. Attend something on The Arts at university (an art exhibition, concert, play or theatre production) and discuss it with students. +0.81 +3.02

Sub-Group: The Sciences (4 items)

45-46 Have serious discussions with students about The Sciences (eg, environment, energy, pollution, genetics, technology or cosmology). +0.77 +3.53

48. Explain an experiment or scientific basis for some aspect of The Sciences (environment, energy pollution, genetics, technology or cosmology). No fit No fit

Sub-Group: Writing (6 items)

50. Prepare an outline in writing to organise the sequence of ideas and points in an assignment. -1.23 +0.68

52. Prepare an outline and write a first draft of a major assignment or project. -0.90 +0.95

53-54 Prepare an outline, discuss it with a peer,

and write and re-write a major assignment. -0.65 +1.75

Sub-Group: Vocations (6 items)

56. Have a serious discussion with students about

Vocations (jobs, occupations, skills). -1.34 +0.93

58. Listen to a lecturer explain or demonstrate

an occupational task or skill, and discuss it

with students. -1.32 +1.11

60. Practice, demonstrate and explain an

occupational task or skill. -0.66 +1.72

Notes

1. The difficulties of the items are in logits.

2. The item difficulties for the Good Coping Strategies, the Actual Coping Strategies and the Coping Strategies Reducing Stress should increase 'horizontally to the right' in correspondence with the conceptual theory of the construct from easy to harder, to harder still.

3. The item difficulties for any one particular response aspect (such as the Actual Coping Strategies) should increase 'vertically down' in correspondence with the conceptual theory of the construct from easy to harder, to harder still.

4. The standard errors vary from 0.08 to 0.10 logits.

5. Items 1-3 use stem-item 1, items 4-6 stem-item 2, and so on up to items 19-21 (stem-item 7).

Appendix B

Item location, standard error, residual and fit to the measurement model

---

Item Stem Location SE Residual DegFree Datapts ChiSq ProbFit

My Course

1 1 <-4 Students scored maximum raw scores. Rasch cannot estimate.

2 1 -3.45 0.93 -0.17 354.76 363 3.29 0.64

3 2 <-4 Students scored maximum raw scores. Rasch cannot estimate.

4 2 -1.71 0.40 -0.59 354.76 363 3.41 0.63

5 3 -2.43 0.57 -1.12 354.76 363 2.98 0.70

6 3 -0.56 0.25 -0.75 354.76 363 3.64 0.59

7 4 -3.63 1.01 -0.71 354.76 363 1.03 0.96

8 4 -0.15 0.21 -0.63 354.76 363 2.86 0.71

9 5 -2.96 0.73 -0.62 354.76 363 1.49 0.91

10 5 -1.70 0.40 -0.67 354.76 363 1.96 0.85

11 6 -1.72 0.41 -1.17 354.76 363 1.30 0.93

12 6 0.63 0.16 0.09 354.76 363 6.99 0.20

13 7 -0.57 0.25 1.15 354.76 363 10.15 0.04

14 7 1.74 0.13 0.09 354.76 363 8.11 0.12

15 8 -1.88 0.44 -0.67 354.76 363 1.77 0.88

16 8 0.35 0.18 -0.99 354.76 363 2.59 0.75

17 9 -1.89 0.44 -0.49 354.76 363 2.53 0.76

18 9 1.38 0.14 1.31 354.76 363 7.59 0.15

The Library

19 11 -2.42 0.56 0.23 354.76 363 7.42 0.17

20 11 0.64 0.16 0.05 354.76 363 3.34 0.64

21 10 -1.08 0.30 -0.65 354.76 363 2.76 0.73

22 10 1.62 0.13 1.38 354.76 363 10.33 0.04

23 12 No fit

24 12 No fit

25 13 -0.97 0.29 -0.13 354.76 363 1.00 0.96

26 13 2.39 0.12 2.34 354.76 363 7.36 0.17

My Lecturers

27 14 -1.81 0.42 0.03 354.76 363 6.68 0.22

28 14 0.81 0.16 -1.63 354.76 363 5.80 0.30

29 15 -1.33 0.34 -0.35 354.76 363 4.19 0.51

30 15 1.90 0.12 -0.23 354.76 363 6.48 0.24

31 16 0.29 0.18 -1.91 354.76 363 5.88 0.30

32 16 2.83 0.12 0.21 354.76 363 5.06 0.39

33 17 0.32 0.18 -0.98 354.76 363 3.39 0.63

34 17 3.83 0.13 0.47 354.76 363 7.20 0.18

Student Acquaintances

35 18 -1.08 0.31 -1.68 354.76 363 4.82 0.42

36 18 0.99 0.15 -0.57 354.76 363 10.21 0.04

37 19 No fit

38 19 No fit

39 20 0.17 0.19 -1.68 354.76 363 4.76 0.43

40 20 2.28 0.12 0.26 354.76 363 3.45 0.62

The Arts

41 21 No fit

42 21 No fit

43 22 0.81 0.16 -1.28 354.76 363 2.24 0.81

44 22 3.02 0.12 0.22 354.76 363 10.30 0.04

The Sciences

45 23 0.77 0.16 0.78 354.76 363 8.58 0.10

46 23 3.53 0.12 1.71 354.76 363 16.17 0.00

47 24 No fit

48 24 No fit

Writing

49 25 -1.23 0.32 -1.95 354.76 363 5.52 0.34

50 25 0.68 0.16 -0.34 354.76 363 0.69 0.98

51 26 -0.99 0.28 -1.35 354.76 363 2.57 0.76

52 26 0.95 0.15 -1.19 354.76 363 2.95 0.70

53 27 -0.65 0.25 -1.91 354.76 363 5.07 0.39

54 27 1.75 0.13 -1.01 354.76 363 2.26 0.81

Voactions

55 28 -1.34 0.34 -1.45 354.76 363 4.62 0.45

56 28 0.93 0.15 -0.20 354.76 363 1.04 0.96

57 29 -1.32 0.34 -1.82 354.76 363 6.57 0.23

58 29 1.11 0.14 -2.03 354.76 363 9.82 0.05

59 30 -0.66 0.26 -1.05 354.76 363 3.08 0.68

60 30 1.72 0.13 -1.23 354.76 363 4.82 0.42

Notes

1. Location (item difficulty) is measured in logits (log odds of answering positively).
2. SE is the standard error in logits.
3. Residual is the item-person interaction test-of-fit statistic for each item based on the degrees of freedom.
4. ChiSq is the item-trait interaction chi-square statistic for each item and ProbFit is the probability of its occurrence for the degrees of freedom listed.

Table I

Summary of global statistics for the Quality of Student Experiences data

(Response categories=2, No. of items=50, No. of students=372)

Item-Student Interaction

Items Students

Location Fit Statistic Location Fit Statistic

Mean +0.00 -0.50 +2.97 +0.37

SD +1.79 +1.00 +1.26 +0.69

Item-Trait Interaction

Total Item Chi Sq 248.09

Total Degree Freedom 250

Total ChiSq Probability 0.52

Student Separation Index is 0.81

Item Separation Index is 0.92

Notes

1. The item-student interaction indicates the degree to which students answer items of different difficulty in a logical and consistent manner. When the data fit the measurement model, the fit statistic has a mean near zero and a SD near 1. A negative fit statistic indicates that the data fit the model very closely. A positive fit statistic indicates that some 'noise' is present.

   The item-student fit statistic relates directly to the consistency of individual student and item response patterns and, hence, to whether there is agreement on the order of item difficulties. In this case it is good, but there is room for improvement.

2. The item-trait interaction indicates the consistency of the item difficulties across the range of different student Quality measures on the scale. When the data fit the measurement model, the item-trait interaction (a chi-square) has a probability greater than 0.01.
3. The power of the tests-of-fit is good, based on the value of the Student Separation Index (proportion of observed variance considered true).

Table II

Expected and actual ranks of item difficulties by Quality Perspective

Vertical Ranking of stem-items

Ideal Perspective Experience Perspective

Item Expected Actual Expected Actual

Stem rank rank rank rank

My Course: Information and Assignments

1 1(easiest) 1 (estimate) 1 1 (very easy)

2 2 2 (estimate) 2 2

3 3(very easy) 3 3 3 (harder, but still easy)

My Course: Discussions

4 1 (very easy) 1 1 2*

5 2 2 2 1* (easiest)

6 3 3 3 3

7 4 (easy) 4 4 4 (hardest)

My Course: Practical

8 1 (easy) 1= 1 1 (hard)

9 2 (harder) 1= 2 2 (harder)

The Library

10 1 (very easy) 1 1 1 (harder)

11 2 2= 2 2 (hard)

12 3 no fit 3 no fit

13 4 (easy) 2= 4 4 (hardest)

My Lecturers

14 1 (very easy) 1 1 1 (hard)

15 2 2 2 2

16 3 3= 3 3

17 4 (hard) 3= 4 4 (hardest)

Student Acquaintances

18 1(easy) 1 1 1 (hard)

19 2 No fit 2 No fit

20 3(hard) 3 3 3 (very hard)

The Arts

21 1 (hard) No fit 1 no fit

22 2 (harder) 2 2 2 (very hard)

The Sciences

23 1 (hard) 1 1 1 (very hard)

24 2 (harder) no fit 2 no fit

Writing

25 1(very easy) 1 1 1 (hard)

26 2 2 2 2

27 3 (easy) 3 3 3 (very hard)

Vocations

28 1(very easy) 1= 1 1 (hard)

29 2 1= 2 2

30 3(easy) 3 3 3 (very hard)

Table III

Expected and actual ranks of item difficulties by Quality Perspective

Horizontal Ranking of Perspectives

Ideal Perspective Experience Perspective

Item Expected Actual Expected Actual

Stem rank rank rank rank

My Course: Information and Assignments

1 1(very easy) 1 2 2 (very easy, but harder)

2 1 1 2 2

3 1(very easy) 1 2 2 (harder, but still easy)

My Course: Discussions

4 1 (very easy) 1 2 2 (easy, but harder)

5 1 1 2 2

6 1 1 2 2

7 1 (easy) 1 2 2 (harder)

My Course: Practical

8 1 (easy) 1 2 2 (hard)

9 1 (easy) 1 2 2 (harder)

The Library

10 1 (very easy) 2 2 2 (hard)

11 1 2 2 2

12 1 no fit 2 no fit

13 1 (easy) 2 2 2 (very hard)

My Lecturers

14 1 (very easy) 1 2 2 (hard)

15 1 1 2 2

16 1 1 2 2

17 1 (hard) 1 2 2 (very hard)

Student Acquaintances

18 1(easy) 1 2 2 (hard)

19 1 No fit 2 No fit

20 1(hard) 1 2 2 (very hard)

The Arts

21 1 No fit 2 no fit

22 1 (hard) 1 2 2 (very hard)

The Sciences

23 1 (hard) 1 2 2 (very hard)

24 1 no fit 2 no fit

Writing

25 1(very easy) 1 2 2 (hard)

26 1 1 2 2

27 1 (easy) 1 2 2 (very hard)

Vocations

28 1(very easy) 1 2 2 (hard)

29 1 1 2 2

30 1(easy) 1 2 2 (very hard)

MEASURES OF QUALITY DIFFICULTIES OF ITEMS

High measures of Quality | Hard items

|

XXXX |

|

6.0 |

|

|

|

XXXXXX |

5.0 |

|

|

XXXXXXXXX |

|

4.0 |

XXXXXXXXXX | Lecturer17.2

|

XXXXXXXXX | Science23.2

XXXXXXXXXXXXX |

3.0 | Arts22.2

XXXXXXXXX | Lecturer16.2

XXXXXXXXX |

XXXXXXXX |

XXXXX | Acquaint20.2 Library13.2

2.0 XXXXXXX |

XXXX | Lecturer15.2

XXXXXXXXX | Library11.2 Vocat30.2 Course7.2 Writing27.2

XX |

XXXX | Course9.2

1.0 XXXX | Vocat29.2

X | Arts22.1 Lecturer14.2 Vocat28.2 Writing26.2 Acquaint18.2

X | Course16.2 Library10.2 Writing25.2 Science23.1

|

X | Lecturer16.1 Lecturer17.1 Course8.2

0.0 | Acquaint20.1

| Course4.2

|

| Course7.1 Course3.2

| Vocat30.1 Writing27.1

-1.0 | Library13.1 Writing26.1

| Acquaint18.1 Library11.1

| Vocat28.1 Lecturer15.1 Vocat29.1 Writing25.1

|

| Course6.1 Course2.2 Course5.2

-2.0 | Course9.1 Course8.1 Lecturer14.1

|

|

| Course3.1 Library10.1

|

-3.0 | Course5.1

|

|

| Course1.2

| Course4.1

-4.0 | <Course1.1 <Course2.1

Low measures of Quality | Easy items

Figure 1. Quality of Student Experiences Scale with student measures and item difficulties calibrated on the same scale.

Notes

1. The scale is in logits and each X represents 3 students. Student scores (left of scale) range from +0.2 to +6.4 logits. Item difficulties (right of the scale) range from -4 to +4.0 logits. The difficulties for items 1 (Course 1.1) and 3 (Course 2.1) couldn't be estimated because every student scored the maximum on these items.
2. The range of item difficulties does not 'cover' the range of Quality scores at the high end of the scale. This means that some harder items have to be added to improve the targeting of the scale.
3. Course refers to stem-items on My Course. Course1.1 is the difficulty for stem-item 1 in the Ideal perspective and Course1.2 is the difficulty in the Experience perspective.
4. Library refers to stem-items on The Library. Library10.1 is the difficulty for stem-item 10 in the Ideal perspective and Library10.2 is the difficulty in the Experience perspective.
5. Lecturer refers to stem-items on My Lecturers. Lecturer14.1 is the difficulty for stem-item 14 in the Ideal perspective and Lecturer14.2 is the difficulty in the Experience perspective.
6. Acquaint refers to stem-items on Student Acquaintances. Acquaint18.1 is the difficulty for stem-item 18 in the Ideal perspective and Acquaint18.2 is the difficulty in the Experience perspective.
7. Arts refers to stem-items on The Arts. Arts21.1 is the difficulty for stem-item 21 in the Ideal perspective and Arts21.2 is the difficulty in the Experience perspective.
8. Science refers to stem-items on The Sciences. Science23.1 is the difficulty for stem-item23 in the Ideal perspective and Science23.2 is the difficulty in the Experience perspective.
9. Writing refers to stem-items on Writing. Writing25.1 is the difficulty for stem-item 25 in the Ideal perspective and Writing25.2 is the difficulty in the Experience perspective.
10. Vocat refers to stem-items on Vocations. Vocat28.1 is the difficulty for stem-item 28 in the Ideal perspective and Vocat28.2 is the difficulty in the Experience perspective.