QUALITY OF STUDENT EXPERIENCES AT UNIVERSITY:

A RASCH MEASUREMENT MODEL ANALYSIS ®

Russell F. Waugh

Edith Cowan University

Paper presented at the Australian Association for Research in Education Conference held in Melbourne, Australia, from 29 November to 2 December, 1999

(WAU99061)

Key words: university, students, expectations, experiences
quality, effort

Running head: QUALITY OF EXPERIENCES

October 1999

Address correspondence to Dr Russell F. Waugh at

Edith Cowan University, Pearson Street, Churchlands,

Western Australia, 6018

ABSTRACT

The Community College Student Experiences Questionnaire (61 items) from the USA (Friedlander, Pace & Lehman, 1990) was revised and rewritten for Australian university students. The Australian Quality of Student Experiences Scale comprises 60 items relating to student expectations and, in direct correspondence, 60 items relating to their experiences. The items are based on a model involving academic, personal and group experiences for eight areas: My Course (18 items), The Library (14), My Lecturers (16), Student Acquaintances (12), The Arts (12), Writing (16), The Sciences (18) and Vocations (14). The convenience sample was 305 1st year students from an Australian university and the data were analyzed with a Rasch measurement model. Fifty-eight items did not fit the model and were discarded. Most of these items came from the sub-scales: The Arts, The Sciences and Writing. The final scale of 62 items had excellent psychometric properties. Expectations are easier than experiences, as conceptualized, and the conceptual design of the scale, as involving My Course, The Library, My Lecturers, Student Acquaintances, Writing and Vocations, is confirmed.
QUALITY OF STUDENT EXPERIENCES AT UNIVERSITY:
A RASCH MEASUREMENT MODEL ANALYSIS

Introduction

In the USA, but not in Australia, there is a great deal of research on college, community college and university experiences that maximize student impact. This research has been reviewed by Pascarella and Terenzini (1991) in their book, *How College Affects Students*. A major finding is that:

'one of the most inescapable and unequivocal conclusions we can make is that the impact of college is largely determined by the individual's quality of effort and level of involvement in both academic and nonacademic activities. This is not particularly surprising; indeed, the positive effects of both the quality and extent of involvement have been repeatedly stressed.'

(Pascarella & Terenzini, 1991, p610)

This means that students who deliberately aim to take part in the varied activities and life of a university or college are more likely to show academic growth, personal growth, satisfaction with their studies and their institution, and continue with their studies. Some examples of the experiences studied include participation in class discussions, use of the library, interaction with lecturers, interaction with students of various religions and groups, discussions and attendance at functions or activities of The Arts (theatre, paintings, sculpture, music recitals, dance, musicals, plays), involvement with writing and literature, using computers, discussions on current and topical science issues, and learning vocational tasks and skills. Students can take part in all these activities which can contribute to a rich and fulfilling experience, if they make the effort.

If this evidence were applied to universities in Australia, there are, at least, two main implications. First, students who show the most academic and personal growth, who are most satisfied with their university and who have a long association with studies at their university, are the ones who put the greatest amounts of effort and time into their university life and studies. Second, individual universities should investigate the links between the university environment and the quality of effort exerted by their students. If they find that students are not making the effort to take part in the varied activities available, then they could develop strategies to improve the situation.

An ordinal level scale to measure quality of effort was developed by Pace (1979a,b, 1984, 1992) and has been used extensively in the USA. It consists of 61 items over eight sub-scales involving My Course, the Library, Faculty Staff, Student Acquaintances, The Arts, Writing, The Sciences and Vocations. Its psychometric properties have been studied by Lehman (1991, 1992) and Ethington and Polizzi (1996), using traditional measurement techniques with USA data. The latter stated that (p711) 'the Quality of Effort measures can be used to make valid and reliable inferences regarding students' efforts and involvement, and that the validity of the inferences is not conditional on whether the students are in vocational or transfer programs, attending full-time or part-time, or of majority or minority ethnic status'.

A scale focusing on Quality of Student Experiences does not seem to have been applied to Australian universities in recent times. Most universities in Australia use the Course Experience Questionnaire (CEQ) which measures student satisfaction with courses and teaching (Johnson, 1997; Waugh, 1999b, 1998; Wilson, Lizzio & Ramsden, 1997). This questionnaire has five sub-scales (25 items in a Likert format): Good Teaching, Clear Goals
and Standards, Appropriate Assessment, Appropriate Workload and Generic Skills, and a single item on Overall Satisfaction with the course. It is based on a model of teaching and is not intended to cover all the other experiences that universities offer.

In 1999, universities first used the Postgraduate Research Experience Questionnaire (PREQ) to investigate postgraduate experiences that were expected to be different from undergraduate experiences. This questionnaire has six sub-scales (28 items in Likert format): Supervision, Thesis Examination, Skills Development, Goals & Standards, Intellectual Climate, Infrastructure and a single item on Overall Satisfaction.

Both the CEQ and PREQ are contentious policy issues in higher education in Australia for three reasons. First, they are used to develop performance indicators (or will be for the PREQ) with comparisons between universities. Second, there is the threat of funding being granted or reduced as a result of performance indicators. Third, there is some disagreement about the validity of the CEQ to measure graduate course experiences four months (or more) after graduation and about whether it samples all the main aspects for all universities. There is even more disagreement about the validity of the PREQ, especially since many supervisors have an insufficient number of postgraduates to satisfy a reliability criterion (and they cannot be identified anyway for ethical reasons) and the intended unit of analysis is broad ‘fields of study’ within a university (rather than supervisors).

Other recent, related Australian studies seem to have focused on the diversity in experiences of university undergraduates (McInnes, James & McNaught, 1995), overseas students and their problems and a comparison with Australian student-problems (see for example, Burke, 1990, 1988a, 1988b, 1986a, 1986b; Mullins, Quintrell & Hancock, 1995; Quintrell, 1992, 1991, 1990). The present study, therefore, has the potential to investigate the psychometric characteristics of an Australian Quality of Experiences Scale for university students and help Australian universities to improve the experiences of their students.

Problems with the USA Quality of Effort Scale

Five aspects of the USA Quality of Effort Scale are called into question. First, if students are asked to respond to items in a 1 to 4 format (from none to very often) and apply this format across all units (subjects) and experiences, then there are problems with interpretation. When students interact with many lecturers and other students, and study many different units (subjects), with different degrees of effort, it is difficult for some of them to answer globally and consistently, as the same amount of time may be interpreted differently by different students. There is a consequent measurement problem for the researcher where the interpretation is unclear. What is needed is an ordered response format that can be applied consistently and logically across all units (subjects) by all students.

Second, the validity of the Quality of Effort Scale is suspect because a proper scale in which the items are ordered from easy to hard has not been constructed. In addition, no attempt has been made to link, on the same scale, the student measures to the item difficulties. Three, the scale only measures student descriptions of their efforts (which influence their experiences during the course). It is likely that their efforts and experiences will be influenced by their expectations (attitudes). Hence, what they expect to experience, as well as what they do experience, ought to be measured at the same time and calibrated on the same scale. That is, many students will have been told what to expect at university (perhaps by teachers, parents and peers) and then by lecturers. Many students will expect university to be a place where a variety of ideas and topics are discussed and debated, not just in relation to a narrow chosen course of study, but in relation to new discoveries, new techniques and the important scientific and artistic 'issues of the day'. Hence, the
expectations of many students will influence their experiences, and be related to their satisfaction with university (see Conceptual Framework).

Four, the main analysis of the Quality of Student Effort Scale has only been performed with traditional measurement techniques and ordinal level scales. That is, the sub-scales are formed from items found to load on various factors through factor analysis and combined to form the final scale. Student scores are formed from just adding the scores on all items (an ordinal level scale) and no check is made to ensure that all items are answered logically as part of a scale (measuring from low to high). Five, its conceptual structure is untested at universities in Australia. Modern measurement programs are now available to create interval level measures in which item difficulties and student measures can be calibrated on the same scale and so test the conceptual structure of Quality of Student Experiences Scale (see Andrich, 1988a, 1988b; Andrich, Sheridan, Lyne & Luo, 1998; Rasch, 1960/1980; Waugh, 1999a,b, 1998).

Changes made to the USA Quality of Effort Scale.

Changes were made to the USA version of the Quality of Effort Scale to overcome the five problems referred to above. The original eight sub-scales (My Course, The Library, My Lecturers Student Acquaintances, The Arts, Writing, The Sciences and Vocations) were retained in the new design. The original 61 items were revised and rewritten so as to be applicable to Australia and written in a positive format. There are now 60 items relating to expectations and, in direct correspondence, 60 items relating to course experiences (see Appendix A). The items were ordered under their respective sub-scale headings which makes it clear to the students what sub-scale is being measured. The response format was then changed in two ways. First, two columns were added for responses, one for expectations and another for experiences. Second, the response categories were changed to an ordered format to provide a better measurement structure: in all the units or nearly all the units (subjects) studied, in most units (subjects) studied, in some units (subjects) studied and in no units or only one unit (subject) studied. The data were analyzed with a Rasch measurement model program to create an interval level scale and to investigate the conceptual structure of the scale (Andrich, Sheridan, Lyne & Luo, 1998).

There are a number of items that may appear to be inappropriate for use in Australia and could have been excluded on simple conceptual grounds, For example, 'I expect to have to explain an experimental procedure to another university student' (Science aspect, items 95-96) and 'I expected to talk about art (painting, sculpture, artists and architecture) with other students at university' (Arts aspect, items 61-62). It is argued that items like these could not be excluded on simple conceptual grounds for two reasons. The first is that we genuinely do expect our best university students to be able to converse on many topics in The Arts and Sciences, no matter what course they are studying, and many can. This is part of the university environment which includes a wide breadth of learning, knowledge and experience. The second is a measurement issue. We know that many students will answer in none or only 1 unit and the scale needs to discriminate these students from the best students. Hence, it is appropriate to include these items to check that they fit the measurement model and that they discriminate between students of differing quality of experiences.

The Australian Quality of Student Experiences Scale places expectations and experiences on the same interval level scale so that a model of expectations and experiences can be constructed from a basis of evidence. Expectations (attitudes) that fit the Rasch model are expected to fall at an easier position on the scale than their corresponding experiences that fit the model. This is done so that expectations and experiences can be compared accurately. It stands in contrast to the usual procedure which is to construct a set of items for
expectations and compare answers on the same set of items for experiences without any statistical or measurement link at the interval level (an invalid procedure).

The new scale uses the terms easy/easier and hard/harder for items relating to expectations and experiences. This may appear, at first reading, to be out of place. While it is common to use easy/difficult to describe achievement items in, for example, a science test, it is not common to use easy/difficult for items relating to expectations and experiences. Nevertheless, this is what is used here and the scale literally indicates that some expectations are easier to hold than their corresponding experiences.

Limitations

There are three main limitations to this study: acceptance of the Rasch model, acceptance of the measurement of expectations and experiences at the same time, and a perceived decrease in validity at the expense of an increase in unidimensionality. With regard to the first limitation, not all measurement researchers accept the Rasch model as valid (see Divgi, 1986; Goldstein, 1980; Traub, 1983). A question arises as to whether the researcher should choose a model that fits the data or use the requirements of measurement to model the data (Andrich, 1989). The Rasch model uses the latter approach. It requires the researcher to define a continuum (from less to more), use a statistical measurement model to check on the consistency of the person measures and the item difficulties, and use a scale in which the property of additivity for item difficulties is valid and the scale values of the statements are not affected by the opinions of people who help to construct it. The former (traditional) approach means that a more complex model is needed to fit the data by increasing the number of parameters to allow for variation in item discrimination and for guessing, as two examples. In that case, the model would then have two person parameters and two item parameters, as a minimum, to model the data.

With regard to the second limitation, the study assumes that students who are surveyed in September and October are able to state reliably their expectations as they were in March. On the surface, this may be questionable and there is evidence that students retrospective recollections can be biased by their implicit theories about personal change (Ross, 1989; Conway & Ross, 1984). It may be that expectations should have been listed first for all items so that all experience items are answered later. However, the questionnaire was trialled with 12 students individually and they were interviewed afterwards. The students said that they were clearly able to separate their expectations from their experiences and, it should be noted, it is expectations at the time of measurement that are related to experiences.

With regard to the third limitation, the Rasch model will reject items that do not fit the model (thus increasing its unidimensionality) but, because there will then be a different number of expectation and experience items, it may be claimed that there is a loss of validity. The counter claim is that there is an increase in validity and unidimensionality. The approach taken in this study is to only use items that contribute to a checkable interval level scale where both expectations and experiences are calibrated on the same scale. The more usual approach is to have a set of expectation (attitude) items and then compare the answers on the same set of items for experiences. This approach is not checking that the data fit a proper measurement scale; it is not checking that expectations and experiences are calibrated on the same scale and hence comparisons of expectations and experiences are then called into question.

Conceptual framework

It is assumed that there is an underlying trait that could be called Quality of Student Experiences at University. This trait would be exhibited as an attitude (expectation) at the
beginning of the course and be modified by experiences during the course. The trait is related to the academic, personal and group experiences for eight aspects associated with student efforts to participate in the life of a university: My Course, The Library, My Lecturers, Student Acquaintances, The Arts, Writing, The Sciences and Vocations. Thus, Quality of Student Experiences is conceptualized, in part, as an expectation derived from eight aspects of university life and, in part, as an experience during university life.

It is conceptualized that, while most students will have high expectations for most of the items in the eight structures for their lives at universities, their experiences will be of a lower standard. That is, they will find most of the items easier in the expectation mode and more difficult in the experience mode. For example, it is conceptualized that students will expect teaching staff to compare and contrast different points of view on many topics (an easy item), but when they come to university they find that many topics are only presented with one point of view (a harder item). Similarly, students will expect library staff to put a lot of time into helping them find articles and material on various topics (easy item), but when they come to university they find that many librarians have only a limited time to spend with each student (hard item). It is theorized that this pattern of easy expectation items, which are harder in experience, will occur for most items and most students, provided the items fit the model and can be placed on the scale. This is in line with the theory that attitudes (expectations) influence behaviour (experiences) (see Ajzen, 1989; Fishbein & Ajzen, 1975; and Waugh, 1999a,b, 1998).

It is conceptualized that students need time in a university environment in order to grow and develop in their knowledge, attitudes and understandings and ultimately, to succeed. Students who make the effort to take part in a variety of experiences that universities offer are more likely to develop a breadth of knowledge, understand varying attitudes and points of view, and bring that knowledge and understanding to solving problems and achieving academic success. As they improve their knowledge and understanding, they will achieve at a higher level. That is, involvement over time in a breadth of experiences offered at university is an indicator of student effort and involvement is a measure of experience that at least partially influences academic success. As student knowledge, understanding and success in academic work grows, students are expected to gain greater acceptance amongst their peers and others, greater self-confidence and greater satisfaction with university.

It is expected that it will require more effort to be involved in some activities and experiences than in others which will, in turn, lead to greater growth, development and understanding. For example, students need to make the effort to compare and contrast a variety of points of view about major issues in a particular course, including an evaluation of the force of particular strengths and weaknesses, rather than just summarize the major points. They need to be able to think logically and apply that logic in a variety of areas. They are more likely to be able to compare and contrast various points and think logically, if they make the effort to be involved in a variety of experiences and if they develop an expectation that this is necessary.

This leads to the view that expectations are related to experiences. Students who expect to take part in the variety of experiences offered at university are more likely than others to make the effort to be involved in a wide variety of experiences and to bring those experiences to bear, where appropriate, in their particular field of study. It is, therefore, expected that high achieving students (like Rhodes Scholars, major prize winners and the best students in university courses) expect to be involved in a variety of areas and display their talent in a variety of areas, not just in one particular subject.
Aims

The present study had three aims. The first was to create an interval level scale for the Quality of Student Experiences Scale. The second was to analyse its psychometric properties using a modern measurement model, the Extended Logistic Model of Rasch (Rasch, 1980/1960, Andrich, 1988a, 1988b) using a modern computer program (Andrich, Sheridan, Lyne & Luo, 1998). The third was to investigate the conceptual design of the Scale and hence contribute to a model of student experiences at university.

Sample and Administration

The convenience sample consisted of 305 1st year students from an Australian university. There are 74 (24.3%) studying in Early Childhood Education, 67 (22.0%) in Business Management, 46 (15.1%) in Biomechanics, 43 (14.1%) Ecology, 42 (13.8%) in Information Technology and Research, 33 (10.8%) in Science.

Following ethics committee approval, the questionnaires were administered at the beginning or end of a lecture, with the permission of the lecturers, towards the end of 2nd semester of 1st year. The purpose of the questionnaire and the study were explained briefly to the students. It was pointed out that Course Expectations and corresponding Course Experiences were required for the eight sub-scales. The questionnaires were anonymous and only grouped data would be reported. Generally, they took 15-20 minutes to complete. Only respondents who supplied complete data sets were used in the study (except for about 12 that had a few missing responses).

Measurement

Seven measurement criteria have been set out by Wright and Masters (1981) for creating a scale that measures a variable. First, each item should be evaluated to see whether it functions as intended. Second, the relative position (difficulty) of each valid item along the scale that is the same for all persons should be estimated. Third, each person's responses should be evaluated to check that they form a valid response pattern. Four, each person's relative score (attitude or achievement) on the scale should be estimated. Five, the person scores and the item scores must fit together on a common scale defined by the items and they must share a constant interval from one end of the scale to the other so that their numerical values mark off the scale in a linear way. Six, the numerical values should be accompanied by standard errors which indicate the precision of the measurements on the scale. Seven, the items should remain similar in their function and meaning from person to person and group to group so that they are seen as stable and useful measures. These criteria are used in creating the Quality of Student Experiences Scale.

Measurement Model

The Extended Logistic Model of Rasch is used with the computer program Rasch Unidimensional Measurement Models (RUMM) (Andrich, Sheridan, Lyne & Luo, 1998) to analyze the data. This model unifies the Thurstone goal of item scaling with extended response categories for items measuring, for example, Course Expectations and Course Experiences, which are applicable to this study. Item difficulties and student measures are placed on the same scale. The Rasch method produces scale-free student measures and sample-free item difficulties (Andrich, 1988b; Wright & Masters, 1982). That is, the differences between pairs of student measures and pairs of item difficulties are expected to be sample independent.
The zero point on the scale does not represent zero expectation or experience. It is an artificial point representing the mean of the item difficulties, calibrated to be zero. It is possible to calibrate a true zero point, if it can be shown that an item represents zero expectation (or experience). There is no true zero point in the present study.

The RUMM program (1998) parameterizes an ordered threshold structure, corresponding with the ordered response categories of the items. The thresholds are boundaries located between the response categories and are related to the change in probability of responses occurring in the two categories separated by the threshold. A special feature of this version of the RUMM program is that the thresholds are re-parameterized to create an ordered set of parameters which are directly related to the Guttman principal components. With four categories, three item parameters are estimated: location or difficulty (d), scale (q) and skewness (h). The location specifies the average difficulty of the item on the measurement continuum. The scale specifies the average spread of the thresholds of an item on the measurement continuum. The scale defines the unit of measurement for the item and, ideally, all items constituting the measure should have the same scale value. The skewness specifies the degree of modality associated with the responses across the item categories.

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 using two 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 estimates for each item and data are combined across all items to give an overall test-of-fit. The latter shows the collective agreement for all items across students of differing measures.

The item-student test-of-fit examines both the response pattern of students across items and 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. The fit statistics approximate a standardized distribution with a mean expectation near zero and a variance near one, when the data fit the model (Wright & Masters, 1982). Negative values indicate a response pattern that fits the model too closely (probably because dependencies are present, see Andrich, 1985) and positive values indicate a poor fit to the model (probably because 'noise' or other measures are present).

Results

The results are set out in two Figures, two Tables and two Appendices. Figure 1 shows the graph of the measures of Quality of Student Experiences for the 305 students and the difficulties of the 120 items on the same scale in logits (the log odds of answering positively). Figure 2 shows the graph of the measures of Quality of Student Experiences for the 305 students and the difficulties of the 62 items that fit the model on the same scale in logits. Table I gives a summary of the Indices of Student Separation (the proportion of observed variance considered true) and fit statistics for the 120 item scale (where 58 items do not fit the model) and the 62 item scale (where all items fit the model). Table II shows a summary of the range and mean item difficulties for the sub-scales of the 62 item scale. Appendix A shows the questionnaire items and the difficulties of the 62 item scale. Appendices B shows, in probability order, the location on the continuum, fit to the model and probability of fit to the model for the 62 item scale.
Psychometric characteristics of the Quality of Student Experiences Scale

The 62 items relating to Quality of Student Experiences have a good fit to the measurement model, indicating a strong agreement between all 305 students to the different locations of the items on the scale (see Table I and Appendix B). That is, there is strong agreement amongst the students to the item difficulties along the scale. The item threshold values are ordered from low to high indicating that the students have answered consistently and logically with the ordered response format. The Index of Student Separability for the 62 item scale is 0.925. This means that the proportion of observed variance considered true is 92.5%. The difficulties of the items have a similar spread along the scale to that of the student measures (see Figure 2). This means that the items are targeted appropriately for the students. The item-trait tests-of-fit indicate that the values of the item difficulties are strongly consistent across the range of student measures. The item-student tests-of-fit (see Table I) indicate that there is good consistency of student and item response patterns. These data indicate that the errors are small and that the power of the tests-of-fit are excellent.

However, there is one problem and this involves the fit of the sub-scales to the model. Thirty-eight out of 46 items for the sub-scales of The Arts (11 out of 12), The Sciences (16 out of 18) and Writing (11 out of 16) did not fit the model. Students did not answer the response categories in a logical and consistent way and they could not agree on the difficulties of the items on the scale. This meant that science students with high quality of experiences, for example, found items from The Sciences easy and other students with high quality experiences found items from The Sciences hard. It was concluded that these 38 items did not contribute to the measurement of the variable with the other items and so these 38 miss-fitting items were discarded. This meant, in effect, that The Arts, and The Sciences were not confirmed as main aspects of the scale of Quality of Student Experiences, at least as measured with these items.

It could be argued that the deletion of items relating to The Arts, The Sciences and Writing means that the model of Quality of Student Experiences has not been fitted to the data properly: that is, there is a reduction in validity. All universities offer The Arts, The Sciences and Writing and many students, but not all, would have experiences relating to these. In the Rasch analysis, the counter claim is that the data must fit the measurement model to be valid and produce a proper scale. The items relating to The Arts, The Sciences and Writing do not fit the measurement model and hence cannot be included in a proper scale (at least as worded for this study). Now it maybe that someone can word some items relating to The Arts, The Sciences and Writing so that all students, irrespective of subject area studied, can agree on their difficulties on a proper scale.

The evidence from the Rasch analysis is that the 62 item Quality of Experiences Scale is valid and reliable. It is suggested that the scale is not context dependent and is relatively sample independent. That is, the scale parameters do not depend on the students who answer the items or on the opinions of the person who constructed the scale. This is a necessary characteristic of a proper scale and is part of the logic of a Rasch model. That all the paired expectation and experience items do not fit the model or that one of the pairs fit the model and the other does not, does not invalidate the scale. On the contrary, only items that fit the model can logically form part of a valid and proper Rasch developed scale.

**Meaning of the Quality of Student Experiences Scale**

The 62 items that make up the variable Quality of Student Experiences are conceptualized as 'my expectation at the beginning of university' and 'my experiences during 1st year university', measured at the same time, from eight main aspects of university life. Only six of
these aspects - My Course, The Library, My Lecturers, Student Acquaintances, Writing and Vocations - are confirmed as contributing to the variable. The 62 items used to measure the main six aspects define the variable (see Appendix A). They have good content validity and they are derived from a conceptual framework based on previous research and theory. While the difficulties of the various items within each aspect vary, their mean values are in order from My Course (easiest), Writing, The Library, Student Acquaintances, My Lecturers, to Vocations (hardest) (see Table II). This, together with the data relating to reliability and fit to the measurement model, is strong evidence for the construct validity of the variable. This means that the students' responses to the 62 items are related sufficiently well to represent the latent variable Quality of Student Experiences at university.

Discussion of the scale

The scale is created at the interval level of measurement with no true zero point of item difficulty or student measure. Equal distances on the scale between measures of Quality of Student Experiences correspond to equal differences between the item difficulties on the scale. Items at the easy end of scale (for example 3,4,5,6,11,12, see Appendix A) are answered in agreement by nearly all the students. Items at the hard end of the scale (for example 42,48,46,64,120, see Appendix A) are only answered in agreement by those students who have high measures of Quality of Student Experiences. In this sample of university students, while most had good Quality of Student Experiences, there were 81 who had less than adequate Quality of Student Experiences (see Figure 2).

The 22 expectation items that fitted the model are mostly, though not all, towards the easy end of the scale (see Appendix A). This means, for example, that the majority of the students found it easy to say that they expected to combine ideas from different sources of information in preparing assignments, that they expected to summarize major points and information from readings and notes and that they expected to participate in class discussions (My Course). They found it easy to say that they expected to ask their lecturers for information about grades, assignments and coursework (My Lecturers), easy to expect to write an outline to organize the sequence of points and ideas in an assignment (Writing) and relatively easy to expect to ask a librarian for help in finding library materials (The Library).

Some of the expectation items are towards the hard end of the scale where students need a high Quality of Student Experiences measure to answer the items positively. For example, students found it difficult to say that they expected to explain material to other students (My Course), difficult to expect to discuss current events, research and university issues with lecturers and difficult to discuss career plans and ambitions with lecturers (My Lecturers). They also, surprisingly, found it difficult to say that they expected to have serious discussions with students of differing political opinions (Student Acquaintances).

The items relating to experiences in My Course are mostly towards the easy end of the scale. Students found it easy to say that they combined ideas from different sources of information in preparing assignments, easy to say that they asked questions in class discussions and easy to say that they participated in class discussions. They also found it relatively easy to say that they did extra readings on topics introduced in classes, that they studied course materials with other students, and that they compared and contrasted different points of view in their course. While students found it easy to say that they used library computers to find books and easy to prepare a list of references using the library, they found it difficult to find interesting material in the library just by browsing.

Most of the items relating to experiences for My Lecturers, Student Acquaintances and Vocations were towards the hard end of the scale. Students found it difficult to say that they have discussed their career plans, ambitions, current events and research issues with their
lecturers (My Lecturers). They found it moderately hard to say that they have had serious discussions with other students of different backgrounds, personal background, philosophy or country (Student Acquaintances). They found it difficult to say that they practiced a vocational task without supervision (or even with a lecturer present), difficult to say that they identified a vocational problem and located information about what to do to solve that problem, or even that they read how to perform an occupational task.

Nearly all the items that fitted the model are easier in their expectation mode than in their corresponding experience mode, as conceptualized. Thus, while students found it easy to expect to participate in class discussions, it was harder to experience this in their courses (My Course); while they found it easy to expect to have to ask a librarian for help in finding information, it was harder to experience; while they found it moderately easy to expect to discuss assignments with their lecturers, it was harder in experience (My Lecturers). While students found it moderately easy to expect to have serious discussions with others from different backgrounds, it was harder in actual experience (Student Acquaintances) and while students found it moderately easy to say that they expected to have to listen to a lecturer explain how to perform an occupational task, it was harder to experience this.

The current Rasch measurement model analysis supports the conceptual design of Quality of Student Experiences as based on a model involving six main aspects. These are My Course, Writing, The Library, My Lecturers, Student Acquaintances and Vocations. In line with this, the analysis supports the view that Quality of Student Experiences can be measured and used as a unidimensional variable based on items in the expectation mode and actual experience mode. This stands in contrast to claims, using traditional measurement techniques, that the Community College Student Experiences Questionnaire (Ethington & Polizzi, 1996; Pascarella & Terenzini, 1991; Pace, 1979a, 1984, 1992), is based on eight main aspects. The current analysis found that most of the items of aspects involving The Arts, Writing and The Sciences did not fit the model. The current analysis supports also the view that Quality of Student Experiences is comprised of an expectation component as well as an actual experiences component, such that the expectation items are easier than their corresponding experience items. While 33 of the 34 experience items from the six main aspects contribute to the scale, only 21 of the corresponding expectation items contribute. This means that the expectation and experience items contribute differently to the scale.

Implications

Ethington and Polizz (1996, p725) claim that there is a ‘strong relationship between the extent to which students become involved in the academic and social systems of educational institutions and their subsequent growth and development and attainment of their educational goals (see Pascarella and Terenzini, 1991, for a review of this research)’. It is suggested that universities should foster student involvement in a variety of educational activities (academic and non-academic) and not just have a focus on the course and assessment. Students learn more, stay longer in education at their university and support their university more when they are involved in its varied activities (Pascarella & Terenzini, 1991). For the University where the current study was undertaken, strategies could be developed to overcome the deficiencies noted below.

There is a question about interpreting the easy and difficult items on the scale in a way that leads to improvements. Students who have Quality of Experience scores at a lower point on the scale than the difficulty of an item have a low chance of answering that item positively. However, this need not necessarily translate to the view that this is a difficult item on which administrators can seek to make an improved action. There are two reasons for this. The first is that the wording of an item may contribute to its difficulty. For example, the item ‘I
expected to discuss my career plans, interests and ambitions with my lecturers’ is difficult because there just isn’t the time or inclination for all lecturers to do this for several hundred of their students. That this is a difficult item doesn’t translate directly into an administrative improvement issue. Thus, if an item is easy, it could still lead to administrative improvements. For example, the item ‘I expected to ask my lecturers for information about grades and assignments’ is easy but is it easy enough? It could be argued that this should be a very much easier item and that this could lead to an administrative action for improvement.

The University, where the current study was performed, has an excellent performing arts centre and its students present many fine performances, a large number of which is free to students and staff. Yet The Arts, as a main area of other student observation and discussion, did not fit the measurement model and a majority of the students in the current study had not seen a play, dance or musical. They did not expect to talk with other students about The Arts and they did not experience The Arts in their 1st year of university.

Similarly, The Sciences, as a main area of student activity, did not fit the measurement model and a majority of students in the current study did not talk with other students about the social, ethical and scientific issues of our times. Many students did not discuss or explain the scientific basis for environmental concerns about such issues as energy, pollution, recycling and genetics. They did not experience having to explain scientific principles and procedures to others. Yet, the University has degree programmes in various science areas and engineering and attempts to promote The Sciences.

In the main areas that did fit the model, there were a number of items that students found to be difficult. These are aspects upon which the University can improve its involvement for the benefit of the university and students. Students found it difficult to agree that they discussed current important issues, assignment comments, career paths, interests, research and university issues with their lecturers. The University could organize more free lectures and seminars on important issues and research to improve student-lecturer interaction and student involvement.

Students found it difficult to agree that they had practised, demonstrated, explained or watched an occupational or vocational task. The University could provide more focus, for a greater variety of students, on links with vocations and occupations, where that is appropriate.

The Library is a central aspect of any university and is one of the focus areas for staff and students. Yet, in the current study, students found it difficult to say that they found some interesting material in the library just by browsing in the stacks and they found it moderately difficult to say that they had asked a librarian for help in finding library materials. Furthermore, the item relating to checking out books from the library to read at home did not fit the model indicating that many students from a range of experiences did not take books home for interest from the university library.

An implication for further research is that a different wording could be tried and analyzed for those items where both expectations and experiences do not fit the model, especially for The Arts, The Sciences and Writing. This is because it may be unsatisfying to have a scale without corresponding expectation and experience items, despite it being technically alright. One may intuitively feel that the content validity is better if both expectation and experience items fit the measurement model for all items.
Conclusion

The Rasch model was useful in constructing a scale of Quality of Student Experiences at university, with items relating to My Course, The Library, My Lecturers, Student Acquaintances, Writing and Vocations. The final scale of 62 items had good psychometric properties. The proportion of observed variance considered true was 92.5%. The threshold values were ordered in correspondence with the ordering of the response categories. The data fitted the measurement model so that the items are arranged from easy to hard on the scale. Item difficulties and student measures are calibrated on the same scale. Items related to The Arts and The Sciences, and many from Writing were rejected since students couldn’t agree on the difficulties or there was misfit to the model. The measurement requirements also meant that different numbers of expectation and experience items fitted the model. Where expectation items and corresponding experience items fitted the model, expectation items were easier than experience items, as conceptualized. The scale is sample independent; the item difficulties do not depend on the sample of university students used or on the opinions of the person who constructed the items. However, the student measures in this study are only relevant to the university involved. Researchers involved with student experiences at university and university administrators should take note of the results of this study.

References


Attitude structure and Function. New Jersey: Lawrence Erlbaum


Mathematical and Theoretical Systems, Elsevier Science

Publishers: Amsterdam (North-Holland).


Burke, B.D. (1988b). *From the airport or an Australian high school: different experiences of overseas students in their first-year at university*. Paper presented to the Asia-Pacific Conference on Student Affairs, organized by the Hong Kong Student Services Association in Hong Kong, July 1988.


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


Programme, Higher Education Division, DETYA.


**APPENDIX A AUSTRALIAN QUALITY OF EXPERIENCES SCALE (Final 62 item scale)**

**QUESTIONNAIRE: EXPECTATIONS AND EXPERIENCES IN UNIVERSITY COURSES**

Please rate the 120 statements, in relation to all the units (subjects) studied in your course, according to the following response format. Place a number corresponding to your *expectation (at the beginning of your course)* and your *experiences (during your course)* on the appropriate line opposite each statement:

In relation to all the units, or nearly all the units (subjects), studied put 3

In relation to most of the units (subjects), studied put 2

In relation to some of the units (subjects), studied put 1

In relation to no units, or only one unit (subject), studied put 0

Example

If your expectation, at the beginning of your course, was to participate in class discussions in all your units, put 3 and, if you only experienced this in some units, put 1.

Item 1/2. I expect to participate in class discussions 3 1

---------------------------------------------------------------

**Item no. Item wording Expectation Experiences**
Sub-Scale: My Course (18 items)

1/2 I expected to participate in class discussions. -0.837 -0.597

3/4 I expected to combine ideas from different sources of information in preparing assignments. -1.317 -1.257

5/6 I expected to summarize major points and information from readings or notes. -1.251 -0.835

7/8 I expected that I would explain material to other students. +0.336 +0.041

9/10 I expected to do additional readings on topics that were introduced in class or lectures. -0.260 -0.228

11/12 I expected to ask questions about points made in class discussions, lectures or readings. -0.889 -0.738

13/14 I expected to study course materials with other students. -0.233 -0.220

15/16 I expected to compare and contrast different points of view presented in my course. -0.267 -0.140

17/18 I expected to consider the accuracy and credibility of information from different sources. -0.099 -0.053

Sub-Scale: The Library (14 items)

19/20 I expected to use the library as a quiet place to read and study materials. No fit No fit

21/22 I expected to read newspapers, magazines and journals located in the library. +0.129 +0.207
23/24 I expected to check out books from the library to 
read at home. No fit No fit
25/26 I expected to use the library computers to find books 
on topics that I wanted. No fit -0.614
27/28 I expected to have to prepare a list of references 
for assignments, using the library. No fit -0.747
29/30 I expected to have to ask a librarian for help in 
finding library materials. -0.177 +0.201
31/32 I expected to find some interesting material in the 
library just by browsing in the stacks. No fit +0.565

**Sub-Scale: My Lecturers (16 items)**

33/34 I expected to ask my lecturers for information about 
grades, assignments and course work. -0.687 -0.400
35/36 I expected to talk (sometimes and briefly) with my 
lecturers after class about course content. -0.207 +0.202
37/38 I expected to make an appointment to meet with my 
lecturers in his or her office, sometimes. No fit +0.828
39/40 I expected to discuss my assignments with my lecturers. -0.182 +0.141
41/42 I expected to discuss my career plans, educational plans 
interests and ambitions with my lecturers. +0.580 +1.170
43/44 I expected to discuss comments made by lecturers 
on assignments that I wrote. No fit +0.489
45/46 I expected to discuss with lecturers (sometimes and 
informally) current events, research and university issues. +0.557 +0.986
47/48 I expected to discuss performance, difficulties or personal 
problems with my lecturers. No fit +1.050
Sub-Scale: Student Acquaintances (12 items)

49/50 I expected to have serious discussions with students who are older and younger than me. No fit -0.055

51/52 I expected to have serious discussions with students whose ethnic or cultural background is different from mine. -0.048 +0.142

53/54 I expected to have serious discussions with students whose philosophy of life & personal values are different from mine. No fit +0.169

55/56 I expected to have serious discussions with students whose political opinions are different from mine. +0.320 No fit

57/58 I expected to have serious discussions with students whose religious beliefs are different from mine. No fit No fit

59/60 I expected to have serious discussions with students from a different country from mine. +0.039 +0.261

Sub-Scale: The Arts (12 items) Expectation Experience

61/62 I expected to talk about art (painting, sculpture, artists and architecture) with other students at university. No fit No fit

63/64 I expected to talk about music (classical, popular) and musicians with other students at university. No fit +0.921

65/66 I expected to talk about theatre (plays, musicals and dance) with other students at university. No fit No fit

67/68 I expected to attend an art exhibition at university. No fit No fit

69/70 I expected to attend a concert or other musical event at university. No fit No fit

71/72 I expected to attend a play, dance concert, or other theatrical performance at university. No fit No fit

Sub-Scale: Writing (16 items)
73/74 I expected to use a dictionary to look up the proper meaning, definition and spelling of words. No fit -0.326

75/76 I expected to prepare an outline to organize the sequence of ideas and points in an assignment. -0.614 -0.535

77/78 I expected to have to think about grammar, sentence structure, paragraphs, and word choice in assignments. No fit No fit

79/80 I expected to have to write a rough draft of an assignment and revise it, before submitting it to my lecturer. No fit No fit

81/82 I expected to use a computer in typing and preparing my assignments. No fit No fit

83/84 I expected to ask other people to read something I wrote to see if it was clear to them. No fit -0.031

85/86 I expected to spend at least 5 hours (or more) writing an assignment. No fit No fit

87/88 I expected to ask my lecturer for advice and help to improve my assignment or to explain comments written on my assignment. No fit +0.135

**Sub-Scale: The Sciences (18 items)**

89/90 I expected to have to memorise formulae, definitions and technical terms. No fit +0.105

91/92 I expected to practice and improve my skills in using laboratory equipment. No fit No fit

93/94 I expected to have to show another university student how to use a piece of scientific equipment. No fit No fit

95/96 I expected to have to explain an experimental procedure to another university student. No fit No fit

97/98 I expected to have to explain my understanding of some
scientific principle by explaining it to other students. No fit No fit

99/100 I expected to complete an experiment or project using scientific methods. No fit No fit

101/102 I expected to talk about social and ethical issues relating science and technology (such as energy, pollution, genetics) No fit +0.795

103/104 I expected to use information learned in a science class to understand some aspect of the world around us. No fit No fit

105/106 I expected to have to explain to someone the scientific basis for environmental concerns about such issues as energy, pollution, recycling and genetics. No fit No fit

**Sub-Scale: Vocations (14 items)**

107/108 I expected to read how to perform an occupational task or vocational skill. No fit +0.569

109/110 I expected to have to listen to a lecturer explain how to perform an occupational task or vocational skill. -0.066 +0.170

111/112 I expected to watch a lecturer demonstrate an occupational task or vocational skill. -0.030 +0.312

113/114 I expected to practice an occupational task or vocational skill monitored by a lecturer or other student. No fit +0.430

115/116 I expected to practice an occupational skill or vocational task without supervision. No fit +0.657

117/118 I expected to identify a vocational problem and locate information about what to do to solve the problem. No fit +0.620

119/120 I expected to diagnose a vocational problem and carry out an appropriate procedure without consultation. No fit +0.813

---------------------------------------------------------------------------------------
Notes

1. The difficulties are in logits (the log odds of successfully answering the item)
2. Negative logit values indicate easy items
3. Positive logit values indicate hard items

APPENDIX B Individual Item Fit (62 items)

<table>
<thead>
<tr>
<th>Label</th>
<th>Location</th>
<th>SE</th>
<th>Residual</th>
<th>ChiSq</th>
<th>Probab</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ex006</td>
<td>I006</td>
<td>-0.835</td>
<td>0.08 0.565 0.539 0.908</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex032</td>
<td>I032</td>
<td>0.565</td>
<td>0.06 0.564 0.545 0.906</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex111</td>
<td>I111</td>
<td>-0.030</td>
<td>0.06 0.350 0.577 0.899</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex109</td>
<td>I109</td>
<td>-0.066</td>
<td>0.06 0.406 0.687 0.872</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex045</td>
<td>I045</td>
<td>0.557</td>
<td>0.06 0.319 0.698 0.870</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex035</td>
<td>I035</td>
<td>-0.207</td>
<td>0.06 0.482 0.701 0.869</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex088</td>
<td>I088</td>
<td>0.135</td>
<td>0.06 0.216 0.739 0.860</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex007</td>
<td>I007</td>
<td>0.336</td>
<td>0.07 0.315 0.761 0.854</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex075</td>
<td>I075</td>
<td>-0.614</td>
<td>0.07 -0.001 1.001 0.795</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex050</td>
<td>I050</td>
<td>-0.055</td>
<td>0.06 0.985 1.038 0.786</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex041</td>
<td>I041</td>
<td>0.580</td>
<td>0.06 0.498 1.131 0.762</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex102</td>
<td>I102</td>
<td>0.795</td>
<td>0.06 0.264 1.165 0.754</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex110</td>
<td>I110</td>
<td>0.170</td>
<td>0.07 0.179 1.178 0.751</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex112</td>
<td>I112</td>
<td>0.312</td>
<td>0.06 -0.078 1.219 0.741</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex048</td>
<td>I048</td>
<td>1.050</td>
<td>0.07 1.218 1.252 0.733</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex039</td>
<td>I039</td>
<td>-0.182</td>
<td>0.06 1.605 1.257 0.731</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex022</td>
<td>I022</td>
<td>0.207</td>
<td>0.06 0.142 1.373 0.703</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ex030</td>
<td>I030</td>
<td>0.201</td>
<td>0.06 1.352 1.391 0.699</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Ex076 I076 0.535 0.07 -0.264 1.544 0.662
Ex008 I008 0.041 0.08 0.625 1.599 0.649
Ex108 I108 0.569 0.06 1.225 1.612 0.646
Ex059 I059 0.039 0.06 -0.060 1.735 0.618
Ex018 I018 0.053 0.07 0.000 1.792 0.605
Ex034 I034 -0.400 0.07 -0.393 1.885 0.584
Ex055 I055 0.320 0.06 0.505 1.956 0.569
Ex033 I033 -0.687 0.07 -0.225 1.995 0.560
Ex011 I011 -0.889 0.08 -0.739 2.215 0.514
Ex021 I021 0.129 0.06 2.168 2.443 0.470
Ex054 I054 0.169 0.06 0.397 2.568 0.447
Ex040 I040 0.141 0.07 -0.375 2.765 0.412
Ex014 I014 -0.220 0.07 0.562 2.770 0.411
Ex010 I010 -0.228 0.07 1.399 3.202 0.342
Ex042 I042 1.170 0.07 -0.418 3.243 0.336
Ex060 I060 0.261 0.06 0.001 3.281 0.330
Ex051 I051 -0.048 0.06 0.253 3.350 0.320
Ex029 I029 -0.177 0.06 1.410 3.453 0.306
Ex116 I116 0.657 0.06 -0.326 3.519 0.297
Ex120 I120 0.813 0.07 -0.492 3.701 0.274
Ex012 I012 -0.738 0.08 -0.442 3.711 0.273
Ex002 I002 -0.597 0.08 0.316 3.716 0.272
Ex009 I009 -0.260 0.06 1.962 3.804 0.261
Ex015 I015 -0.267 0.07 -0.558 3.982 0.241
Ex003 I003 -1.317 0.09 -0.021 4.027 0.236
Ex004 I004 -1.257 0.09 0.987 4.028 0.236
Ex013 I013 -0.233 0.07 1.495 4.109 0.227
Ex114 I114 0.430 0.06 0.030 4.121 0.226
Ex017 I017 -0.099 0.07 0.755 4.397 0.198
Ex052 I052 0.142 0.06 0.374 4.728 0.168
Ex038 I038 0.828 0.07 -0.391 4.739 0.167
Ex005 I005 -1.251 0.09 1.157 4.921 0.152
Ex044 I044 0.489 0.06 -0.558 4.965 0.149
Ex016 I016 -0.140 0.08 -0.048 4.984 0.148
Ex036 I036 0.202 0.07 -0.230 5.377 0.120
Ex001 I001 -0.837 0.07 0.811 5.759 0.097
Ex046 I046 0.986 0.07 -0.582 5.811 0.094
Ex028 I028 -0.747 0.07 0.957 5.920 0.088
Ex084 I084 -0.031 0.06 1.733 7.288 0.034
Ex118 I118 0.620 0.06 -1.114 9.043 0.000
Ex090 I090 0.105 0.06 3.272 11.703 0.000
Ex064 I064 0.921 0.07 2.252 13.587 0.000
Ex074 I074 -0.326 0.06 1.287 14.656 0.000
Ex026 I026 -0.614 0.07 1.697 15.513 0.000

1. location is the item difficulty on the scale
2. Residual is the observed response minus the expected value
3. Probab is the chi-square probability fit to the model.

It is sensitive to sample size and should not be interpreted too strictly
Figure 1: Measures of Quality of Student Experiences and item difficulties for the 120 item scale in logits (the log odds of answering positively).

Figure 2: Measures of Quality of Student Experiences and item difficulties for the 62 item scale in logits (the log odds of answering positively).
Table I

Summary data of the reliabilities and fit statistics to the model for the 120 item and 62 item scales (N = 305)

120 item scale 62 item scale

Non-fitting items 58 none
Disordered thresholds 58 none
Index of Student Separability n/a 0.925
Item-trait interaction (chi-square) 548 (p<0.001) 223 (p<0.05)
Item fit statistic mean +0.336 +0.480
Sd +1.005 +0.848
Student fit statistic mean -0.232 -0.171
Sd +2.577 +2.092
Power of test-of-fit n/a excellent

Notes on Table I

1. The Index of Student Separation is the proportion of observed variance that is considered true.
2. The item and student fit statistics have an expectation of a mean near zero and a standard deviation near one, when the data fit the model.
3. The item-trait interaction test is a chi-square. The results indicate that there is good collective agreement for all items across students of differing Quality of Student Experiences.

Table II

Range and mean scores for the sub-scales in the 62 item scale

Sub-Scale Highest Score Lowest Score Mean Score
My Course (n=18/18) +0.336 -1.317 -0.491
Writing (n=5/16) +0.135 -0.614 -0.274
The Library (n=7/14) +0.565 -0.747 -0.062
Student Acquaintances(n=7/12) +0.261 -0.055 +0.118
My Lecturers (n=13/16) +1.170 -0.687 +0.348
Vocations (n=9/14) +0.813 -0.066 +0.386

The Arts (n=1/12) n/a n/a n/a

The Sciences (n=2/18) n/a n/a n/a

Notes on Table II

1. Numbers in brackets represent items fitting the model out of the total possible

2. Table scores are in logits