A Rasch Measurement Model Analysis of the Revised Approaches to Studying Inventory

Russell F. Waugh
Edith Cowan University

and

Patricia A. Addison
Curtin University of Technology

Running head: APPROACHES TO STUDYING

June 1997

Address correspondence to Dr Russell Waugh at
Edith Cowan University, Pearson Street, Churchlands, Western Australia, 6018.
Background. The Revised Approaches to Studying Inventory (Entwistle & Tait, 1994) comprises 38 self-report items designed to measure student approaches to learning in a higher education context. The items have been conceptualised and designed from five learning orientations, corresponding to five sub-scales of the Inventory: ‘a deep approach’, ‘a surface approach’, ‘a strategic approach’, ‘a lack of direction’ and ‘academic self-confidence’.

Aims. The study aims to create an interval level scale for the Inventory and analyse its psychometric properties using a modern measurement model, the Extended Logistic Model of Rasch (Andrich, 1988a, 1988b; Rasch, 1980), and investigate the conceptual design of the Inventory.

Sample. The sample was 346 students (170 females, 176 males, 212 less than 23 years and 134 older than 23), studying in first year Faculty of Business, at a university in Perth, Western Australia.

Method. A scale was created for the Inventory and analysed for reliability, fit to the model, meaning and validity. The inventory was analysed separately for each of four sub-groups (females, males, younger and older students) to test the invariance of the scale. The five sub-scales were analysed separately to test the conceptual design and meaning of the inventory.

Results. The Inventory has satisfactory psychometric properties, when items 20 and 33 are deleted. Easier and harder items need to be added to better target the student measures. Its conceptual design from the five learning orientations is confirmed. Only 15 items have satisfactory invariance across the four sub-groups. The psychometric properties of three sub-scales (deep approach, surface approach and strategic approach) are only moderately satisfactory and the item separabilities of the other two (lack of direction and academic self-confidence) are unsatisfactory.

Conclusions. The Extended Logistic Model of Rasch was found to be useful in creating an interval level scale for the Inventory, and for analysing its psychometric properties and conceptual design. It is suggested that the Inventory can be improved conceptually by adding more items relating to attitudes, intentions and behaviour and by adding harder and easier items to target the student measures better.

Background

The inventory contains 10 items relating to ‘a deep approach’ to studying (looking for meaning, active and critical stance, relating and organising ideas, and using evidence and logic), 10 items relating to ‘a surface approach’ to studying (relying on memorisation, difficulty in making sense, unrelatedness, and concerns about coping), 10 items relating to ‘a strategic approach to learning’ (determination to excel, effort in studying, organised studying, and time management), 4 items relating to ‘a lack of direction’ in studying and 4 items relating to ‘academic self-confidence’ in studying (see the Appendix). There are 5 response categories per item: agree, agree somewhat, unsure, disagree somewhat, and disagree. While there is no denying that the inventory provides good data that are needed by universities and researchers, there are three aspects of the inventory which are called into question.

First, the Revised Approaches to Studying Inventory only provides data at the nominal or ordinal (ranking) level, not at the interval level. A modern measurement model (such as
Rasch) could be applied to the data to construct an interval level measure. Second, a Rasch analysis could help to explain the conceptual structure of the inventory (that is, its meaning and validity) and test whether it is targeted correctly (that is, if the student measures and the item difficulties correspond on the same scale). Third, the neutral response category allows a variety of answers such as don’t know, not sure, neutral or don’t want to answer and this is not good measurement procedure because interpretation is unclear (see Andrich, de Jong & Sheridan, 1994; Dubois & Burns, 1975; Sheridan, 1993). These questions can be answered with the present study.

There are four other aspects which are called into question, but not answered by the present study. First, the inventory doesn’t include items focusing on all the major issues about how students study and learn. For example, it doesn’t include items on ‘no real approach’, discussions of difficult topics and issues with others, interest in the topic or study area, and feelings towards study, including health and environment problems. Second, there is no clear focus directed either towards different items relating to attitudes, intentions and actual behaviour (see Ajzen, 1989), or towards the main issues in health and environmental problems relating to studying (Price, Dunn & Dunn, 1991). That is, the structure of the concept, approaches to learning and studying, is called into question. Third, while the response categories (agree, agree somewhat, unsure, disagree somewhat and disagree) are commonly used in attitude scales, they are not ordered from low to high. In modern measurement, an example of a better response structure is ‘never’ to ‘all-the-time’. Four, while negatively and positively worded items are used to overcome the fixed response syndrome (a common procedure in traditional measurement), the mixture can cause an interaction effect between items in modern measurement models (see Andrich & van Schouwbroeck, 1989; Sheridan, 1995). For modern measurement models, it is often considered better to word all items in a positive sense.

Outline of the Literature

The study of how university students approach learning and studying is an important topic for lecturers, student advisors and counsellors. All of the research on this topic involves traditional measurement methods of scale construction using, for example, factor analysis to determine the orientations, correlations for validity and Cronbach alpha for reliability. No studies of this topic could be found that used a modern measurement model such as Rasch.

Our review begins with Martin & Saljo (1976a,1976b) who found that Scandinavian students used either a deep approach (looking for meaning, questioning, relating to other topics and experience) or a surface approach (focusing on memorising). Following this line, various studies continued on and around this topic and various versions of student self-report measures were researched and tested (see for example Bowden, 1986; Clarke, 1986; Entwistle, Hanley & Hounsell, 1979; Entwistle & Ramsden, 1983; Entwistle, 1988; Harper & Kember, 1989; Murray-Harvey, 1994; Newstead, 1992; Richardson, 1995, 1994, 1993, 1990). This research complemented independent work carried out by Biggs (1987, 1985, 1979) who derived a three factor model of the learning process involving a surface approach, a deep approach and an achieving approach. A recent outcome of all this research was a revised scale, the Revised Approaches to Studying Inventory (38 items) containing five factors - ‘a deep approach’ with 10 items, ‘a surface approach’ with 10 items, ‘a strategic approach’ with 10 items, ‘a lack of direction’ with 4 items and ‘academic self-confidence’ with 4 items (Entwistle & Tait, 1994).

Research showed that students, identified as using ‘a deep approach’, report that they try to work out the meaning of information; they do not accept ideas without some critical evaluation; they relate ideas from their studies to a wider context; and they look for reasoning, justification and logic behind the ideas. Students, identified as using ‘a surface
approach', report that they rely on rote learning; they accept ideas without necessarily understanding them; they emphasise the acquisition of factual material; and express anxiety in terms of the organisation and volume of material. In contrast, students, identified as using 'a strategic approach', perceive themselves as having clear goals related to their studies and being hard workers; they ensure that they have the appropriate resources and conditions; and they are well organised. The fourth orientation to learning reflects the lack of clear, academic and career direction and goals of some students, in contrast to others. The fifth orientation to learning reflects the perception of some students, in contrast to others, as able and intelligent and easily capable of coping with the intellectual and academic demands of their studies.

Aims of the Study

This study aims to create an interval level measure of the Revised Approaches to Studying Inventory (Entwistle & Tait, 1994), using the Extended Logistic Model of Rasch (Andrich, 1988a, 1988b; Rasch, 1960/1980), and investigate its psychometric properties. A further aim is to investigate the conceptual design of the Inventory.

METHOD

Data

A sample of 346 first year students from the Faculty of Business at a University in Western Australia volunteered to answer the questionnaire in 1996. The students came from various departments such as accounting, management, business, finance and information technology. There were 170 females, 176 males, 212 students less than 23 years of age and 134 students older than 23. The original questionnaire response format was reduced from five to four categories for two reasons. First, unsure was omitted because it is not clear what an answer means (see the background comments earlier). Second, an ordered response format from high to low was needed (agree, agree somewhat, disagree somewhat and disagree).

Measurement

Taken individually, the 38 items of the Revised Approaches to Studying Inventory can be used to interpret the responses of students to their approaches to study at university. This could provide a view of their approach to studying from a qualitative point of view on each item. However, if data on the 38 items are aggregated in some way or used to create a scale and then interpreted, then seven criteria have to be met before it can be said that the items form a valid and reliable scale.

The seven measurement criteria have been set out by Wright & Masters (1981). They involve, first, an evaluation of whether each item functions as intended; second, an estimation of the relative position (difficulty) of each valid item along the scale; third, an evaluation of whether each person’s responses form a valid response pattern; four, an estimation of each person’s relative score (attitude or achievement) on the scale; 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; and 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. The present study used these seven criteria to analyse the 38 items of the Revised Approaches to Studying Inventory and its five sub-scales.
Measurement Model

The Extended Logistic Model of Rasch (Andrich, 1978, 1988a, 1988b; Rasch, 1980; Wright, 1985) was used with the computer program Quest (Adams & Khoo, 1994) to create a scale, satisfying the seven measurement criteria of Wright & Masters (1981) and based on the 38 items of Revised Approaches to Studying Inventory. The scale is based on the log odds (called logits) of students agreeing with the items. The items are ordered along the scale at interval measurement level from easiest with which to agree to hardest with which to agree. Items at the easiest end of the scale (those with negative logit values) are answered in agreement by most students and those items at the hardest end of the scale (those with positive logit values) are most likely to be answered in agreement only by students whose attitudes are strongly positive. 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 item difficulties are expected to be sample independent.

The program checks on the consistency of the student responses and calculates the scale score needed for a 50 per cent chance of passing from one response category to the next; for example, from agree to agree somewhat, from agree somewhat to disagree somewhat, and from disagree somewhat to agree for each item. These scale scores are called threshold values. They are calculated in logits and they must be ordered to represent the decreasing attitude needed to answer in order from agree to agree somewhat, from agree somewhat to disagree somewhat, and from disagree somewhat to disagree for each item. Items whose thresholds are not ordered - that is, items for which the students do not use the categories consistently - are not considered to fit the model and are discarded.

The program checks that the student responses fit the measurement model according to strict criteria. The criteria are described by Adams & Khoo (1994), Wright & Masters (1982) and Wright (1985). The fit statistics are weighted and unweighted mean squares that can be approximately normalised using the Wilson-Hilferty transformation. The normalised statistics are called infit t and outfit t and they have a mean near zero and a standard deviation near one when the data conform to the measurement model. A fit mean square of 1 plus x indicates 100x% more variation between the observed and predicted response patterns than would be expected if the data and the model were compatible. Similarly, a fit mean square of 1 minus x indicates 100x% less variation between the observed and predicted response patterns than would be expected if the data and the model were compatible. In this study of the Revised Approaches to Studying Inventory, each item had to fit the model within a 30% variation between the observed and expected response pattern or it was discarded. While the 30% is somewhat arbitrary, it does represent a balance between too much item dependency (where all the items might be too similar) and too much item independence (where each item measures a different aspect).

Reliability is calculated by the Item Separation Index and the Student Separation Index. Separation indices represent the proportion of observed variance considered to be true. A value of 1 represents high separability in which errors are low and item difficulties and student measures are well separated along the scale (Wright & Masters, 1982). A combination of data is required as evidence for the construct validity of the Inventory. The Item and Student Separation Indices need to be high; the observed and expected item response patterns need to fit the measurement model according to strict criteria; the thresholds relating to passing from one category response to the next need to be ordered; and there needs to be a conceptual framework (theoretical or practical) linking the items of the scale together.
Data Analysis

The data were analysed initially with the whole sample (N=346) and all the 38 items together. Items not fitting the model were then discarded (items 20 and 33) and the analysis repeated. This was done to test the meaning, targeting, validity and reliability of the inventory.

The analysis was repeated with each of the four sub-groups of the sample (N=170 females, N=176 males, N=212 less than 23 years, and N=134 older than 23 years). This was done to test that the inventory is used consistently by females, males, younger and older students. It is part of the good measurement criteria set out by Wright & Masters (1981).

The analysis was repeated with each of the five sub-scales separately (deep approach, surface approach, strategic approach, lack of direction and academic self-confidence). As the Inventory was designed and conceptualised from five orientations of learning and studying, the items of each orientation (that is, each sub-scale) should form a satisfactory scale by itself, as well as forming a satisfactory Inventory (with all 38 items together). That is, the Inventory, as a measure of a unidimensional underlying trait called orientations to learning and studying, is conceptualised as overarching the five individual orientations. The separate analysis of the sub-scales from the Inventory is part of analysing the meaning of the Inventory and the trait it measures.

Results

In the interest of conciseness and brevity, not all the results are presented here; only those that are considered the most important. For example, the values for the standard errors of measurement for each item difficulty and student measure are not presented; and the threshold values for each response category of each item for each of the scales created are not included.

The results are presented in three Tables and three Figures Table 1 gives a summary of the scale statistics for the five sub-groups (females, males, younger students, older students and all students). Table 2 gives a summary of the scale statistics for the five sub-scales (deep approach, surface approach, strategic approach, lack of direction and academic self-confidence). Table 3 gives the difficulties of the items by sub-groups. Figure 1 gives the scale for the 36 items with item difficulties and the student measures calibrated on the same scale for the whole sample. Figure 2 gives the fit map for the 36 items fitting the model. The original 38 items of the Inventory (Entwistle & Tait, 1994) are given in the Appendix.

DISCUSSION

Psychometric Characteristics of the Revised Approaches to the Studying Inventory

The final 36 items of the Inventory (with items 20 and 33 deleted) have a good fit to the measurement model, indicating a strong agreement between all 346 students located at different positions on the scale, across all 36 items (see Figures 1 & 2). However, the items
are not as well targeted against the student measures in this sample (1st year university) as they could be. Some harder and easier items are needed (see Figure 1). The item threshold values are ordered from high to low indicating that the students have answered consistently with the ordered response format of agree, agree somewhat, disagree somewhat, and disagree. The Indices of Student and Item Separation are 0.86 and 0.79 (see Table 1) indicating that the separability of the scale is satisfactory (although this could be improved) and that the power of the tests of fit to the measurement model is satisfactory. The values of the infit mean squares and outfit mean squares are approximately 1, the values of the infit t-scores and outfit t-scores are approximately zero (see Table 1) and the mean squares are within 30% of the expected values, calculated according to the model (see Figure 2). This indicates that there is a good fit to the measurement model. The item difficulties and the student measures are calibrated on the same scale with reasonably small errors. However, the difficulties of the items are not invariant with respect to gender and age (see Table 3). That is, the difficulties of some of the items vary too much from group to group and only 15 items have difficulties which could be considered invariant across the groups, within the measurement error. Consequently, an important aspect of creating a scale (sample-free item difficulties) has not been completely realised. It cannot therefore be claimed that the 36 items of the Inventory has completely satisfactory psychometric properties.

---

Place Table 1 about here.

---

**Meaning of the Revised Approaches to Studying Inventory**

The 36 items that make up the variable Student Approaches to Studying relate to five learning orientations. They are a deep approach, a surface approach, a strategic approach, a lack of direction and academic self-confidence. The 36 items define the variable. They have good content validity and they are derived from a conceptual framework based on previous research and theory. This, together with the previous 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 36 items are related sufficiently well to represent the unobservable trait of student achievement in approaches to studying.

---

Place Table 2 about here.

---

**Overview of the scale**

The scale for the Inventory is created at the interval level of measurement with no true zero point of item difficulty or student measure. Equal distances between measures of approach to studying correspond to equal differences between the difficulties of the items on the scale. Items at the easy end of the scale (numbers 2, 30 and 35, for example, see Appendix and Figure 1) are answered in agreement by nearly all the students. Items at the hard end of the scale (numbers 3, 8, 17 and 26, for example) are only answered in agreement by those students who have high achievement in approaches to studying.

The wording of the ten items of the surface approach to studying indicate that they mainly relate to behaviours about studying. They fall at the hard end of the scale. This means that
students found it hard to agree with (achieve) these items (interpreted in a negative sense). That is, for example, students found it hard to agree that they are not drowning in the sheer amount of material with which they have to cope, hard to agree that they don't have to memorise a good deal of what they have to learn, hard to agree that they don't have to worry that they will be able to cope, and hard to agree that they don't find themselves reading without understanding. Students need to have a high achievement in approach to studying to be able to agree with the negative surface approach items.

Eight of the ten items of the deep approach to studying fall at the easy end of the scale. The wording indicates that they mainly relate to trying to study in certain ways; ways that are sensible but difficult to attain. This means that students found it easy to agree with these items in attitude. That is, for example, students found it easy to agree that they set out to understand the meaning of what they have to learn, easy to agree that it’s important to follow the argument and see the reasoning, easy to agree that they try to work out for themselves what’s being said, and easy to agree that they try to relate ideas to other topics. Students only need a weak achievement in approach to studying to agree with these attitude items and most students did agree.

Two of the ten items of strategic approach to studying are the easiest items on the scale. This means that students found it very easy to obtain their text books and easy to feel that they are doing as well as they can. The other eight items are spread along the scale in difficulty from easy to moderately hard. While the wording of most of these items indicates that they refer to attitudes, some of the items refer to behaviours such as working steadily, getting the right conditions for studying and keeping one’s mind on studying.

The eight items relating to lack of direction and academic self-confidence range from moderately easy to moderately hard on the scale. The wording of the items indicates that they relate to attitudes.

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

Place Table 3 about here.

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

Comments on the design of the scale

The analysis supports the conceptual design of the Inventory. That is, it supports the view that the underlying trait, approaches to studying, is an overarching concept comprised of five orientations to studying (deep approach, surface approach, strategic approach, lack of direction and academic self-confidence).

The wording of the items of the sub-groups, deep approach and strategic approach, seem mainly to relate to attitudes, while those of the sub-group, surface approach, seem mainly to relate to behaviours. The items of the other two sub-groups, lack of direction and academic self-confidence, seem mainly to relate to attitudes. In principle, there is no problem including both attitudinal and behavioural items from these five sub-groups on the same scale, since they can be shown to be linked in theory and in measurement. However, because there are no behavioural items from four of the sub-scales and no attitudinal items from the surface approach sub-scale, the analysis gives what appears to be strange results. In the main, the measurement model places the surface behavioural items as the most difficult, then mixes attitudinal items from the other sub-groups in the middle of the difficulty range, and then places attitudinal items from the deep and strategic approach as the easiest.
If behavioural items relating to the deep and strategic approach were added to the scale, it would theoretically be expected that they would fall at the difficult end of the scale, and be more difficult than the surface approach behavioural items. If attitudinal items relating to the surface approach were added to the scale, these would theoretically fall at the easy end of the scale, but be more difficult than the attitudinal deep approach items. That is, it would theoretically be expected that students would find it easy to hold positive beliefs about approaches to studying, harder to hold positive attitudes and even harder still to behave positively, in relation to approaches to studying. The Extended Logistic Model of Rasch (Andrich, 1988a) is therefore of considerable help here because it points out the theoretical problems with the Inventory.

---

Place Table 2 about here.

---

**Items Not Fitting the Model for the Inventory**

Two items, numbers 20 and 33, did not fit the model for the Inventory. Both of these items are from the surface sub-scale and both were reversed scored. Item 20 includes two aspects, repeating and copying to help with memorising, and it seems that the two aspects are causing conflicting responses. It is suggesting that these two aspects be separated into two items. Item 33 includes the two aspects panic and behind with work. Some students apparently panic whether they are behind with work or not and this causes conflicting responses. It is suggested that the wording be modified to remove this conflict.

**The Sub-Scales**

**The Deep Approach Sub-Scale**

The ten items that make up the sub-scale, deep approach, relate to looking for meaning, relating and organising ideas, and using evidence and logic. The ten items have a good fit to the measurement model and ordered thresholds indicating that the responses are answered consistently across the four categories. However, the student and item separabilities are too low (+0.45 and +0.76, see Table 2) and the difficulties of the items are not as well targeted against the student measures as they could be (scale not included here). It would seem that the Deep Approach Sub-Scale has some satisfactory psychometric properties, but it needs revision. The scale could be improved by adding more items to obtain a better targeting against the student measures (see Figure 1) and this could include behavioural items in relation to a revised conceptual design of the inventory.

**The Surface Approach Sub-Scale**

The ten items that make up the sub-scale, surface approach, relate to relying on memorisation, difficulty in making sense, unrelatedness, and concern about coping. Items 7, 20 and 36 did not fit the measurement model and so they would need to be modified. The other seven items have a good fit to the measurement model. All the ten items have ordered thresholds indicating that the responses are answered consistently and moderate student and item separability indices from +0.73 and +0.77 (see Table 2). The difficulties of the items are not as well targeted against the student measures as they could be because there are too few middle-difficulty items (scale not included here). While the surface approach sub-scale has some good psychometric properties, it needs modification. It is suggested that more items be added in relation to surface attitudes.
The Strategic Approach Sub-Scale

The ten items that make up the sub-scale, strategic approach, relate to determination to excel, effort in studying, organisation in studying, and time management. Item 2 did not fit the measurement model and would need to be modified. The other nine items have a good fit to the measurement model. All the ten items have ordered thresholds indicating that the responses are answered consistently and moderate student and item separability indices of +0.73 and +0.80 (see Table 2). The difficulties of the items are not as well targeted against the students as they could be because there are too few middle-difficulty and difficult items (scale not included here). While the strategic approach sub-scale has some good psychometric properties, it needs modification. It is suggested that more items be added, especially in relation to strategic behaviours, in order to make the inventory conform better to its conceptual design and to target the items better.

The Lack of Direction Sub-Scale

The four items of the sub-scale, lack of direction, relate to drifting into the course, pleasing others, unsure of choice, and working to the expectation of others. The four items just fit the measurement model within 33% of their observed and predicted response patterns. However, while they have ordered thresholds indicating that the responses are answered consistently, they also have low student and item separability indices of +0.17 to +0.51 (see Table 2). The low item separability is directly attributable to the low number of items, the poor targeting of the items at the students (data not shown here) and probably to the general focus of the items. More items, including behavioural items, could be added, to provide better targeting and improve the psychometric properties of the sub-scale.

The Academic Self-Confidence Sub-Scale

The four items of the sub-scale, academic self-confidence, relate to student perceptions of their understanding of their subjects, whether they find the work easy, difficulty in making sense of new information and ideas. The four items have a reasonable fit to the measurement model, ordered thresholds indicating that the responses are answered consistently, but moderate or low separability indices of +0.78 to +0.49 (see Table 2). The low item separability is directly attributable to the low number of items and the poor targeting of the items at the students. More items, including behavioural items, could be added to provide better targeting and improve the psychometric properties of the sub-scale.

Conclusion

The Extended Logistic Model of Rasch was useful in creating a good interval level measure of the Revised Approaches to Studying Inventory devised by Entwistle & Tait (1994) and for investigating its validity and reliability. The Rasch model was also helpful in analysing the conceptual design of the Inventory. As a result of this analysis, the Inventory is confirmed as a measure of an overarching trait based on five orientations to learning. The analysis also leads to suggestions for improving the targeting of items against the student measures through the addition of new items relating to attitudes, intentions and behaviour and by modifying some of the less satisfactory items.
References


