Creating a scale to measure Motivation to achieve academically:

Linking attitudes and behaviours using Rasch measurement

A refereed full-paper presented at the Australian Association for Research in Education

Conference held in Fremantle from Monday 3 to Thursday 6 December 2001

Russell F. Waugh

Edith Cowan University

Key words: Motivation, achievement, students, university attitudes, behaviour, measurement, scale, Rasch

Running head: ACADEMIC MOTIVATION

3 July 2001

Address correspondence to Dr Russell F. Waugh at

Edith Cowan University, Pearson Street, Churchlands,

Western Australia, 6018

Abstract

Background. Motivation to Achieve Academically has been used in many educational and other studies in many countries and the large majority has not used an interval level scale based on a good theoretical model in which the items are linked to behaviour.

Aims. One, to create an interval level, unidimensional scale of Motivation, with attitude items linked to behaviour items, based on a conceptual model of Motivation, involving Striving for Excellence (Standards, Goals, Tasks, Effort, Values and Ability), Desire to Learn (Interest, Learning from Others and Responsibility for Learning), and Rewards (Extrinsic, Intrinsic and Social). Two, to analyse its psychometric properties using the Extended Logistic Model of Rasch (Andrich, 1988a, 1988b; Rasch, 1980/1960). Three, to investigate the structure and meaning of the Scale.

Sample. The convenience person sample was 239 first-year students selected for three special entry programs at an Australian university. The stem-item sample was initially 45, reduced to 24, that fitted the measurement model to form a valid and reliable scale.

Method. Based on recent literature, a conceptual model of Motivation was devised and items written in line with the model. Data were collected by self-report questionnaire and analysed with the computer program Rasch Unidimensional Measurement Models (RUMM) (Andrich, Lyne, Sheridan, & Luo, 1998). A scale was created in which the Motivation measures were calibrated on the same scale as the item 'difficulties'.

Results. Twenty-four Motivation items fitted the model and were 'easier' than their corresponding behaviour items, as conceptualised. They formed an excellent scale in which the proportion of observed variance considered true was 0.93. Items from all aspects of the Motivation model named in the aims above, except Ability and Extrinsic Rewards, fitted the measurement model.

Conclusion. The Rasch model and the RUMM computing program were very useful in creating a unidimensional, interval level scale of Motivation to achieve academically, with good psychometric properties.

Creating a scale to measure Motivation to achieve academically: Linking attitudes and behaviours using Rasch measurement

Motivation is a widely used variable in many educational and other studies. While it is difficult to define in words, the following simple description captures its essence.

Motivation involves the internal processes that give behaviour its energy and direction. Motivation originates from a variety of sources (needs, cognitions and emotions) and these internal processes energize behaviour in multiple ways such as starting, sustaining, intensifying, focusing, and stopping it.

(Reeve, 1996, p2)

While it is claimed that Motivation is composed of a number of dimensions, there is disagreement about the specific composition of those dimensions (Donohue & Wong, 1997; see also the special section on Motivation and Efficacy in the Journal of Educational Psychology, 82,1, (Special Editor, Schunk, 1990). Motivation is also claimed to be the

product of interdependence between and amongst many variables (Weiner, 1990), such as locus of control (Duke & Nowicki, 1974), the need for affiliation, impulsiveness and planfulness (Friis & Knox, 1972), personal achievement, social achievement, academic achievement (Maehr, 1984; Piedmont, 1989), mastery, work orientation, competitiveness and personal concern (Helmreich & Spence, 1978; Donohue & Wong, 1997). A large amount of other research has been done on these and related issues of Motivation (see Ames & Ames, 1989, 1985, 1984; Anderman & Maehr, 1994; Brophy, 1987; Covington, 1992; Graham, 1994; Jacobs & Newstead, 2000; Maehr & Ames, 1989; Maehr and Pintrich, 1997, 1995, 1993; Zimmerman, Bandura & Martinez-Pons, 1992).

There are at least ten models of Motivation in the literature, each emphasizing different aspects, some of which are inter-related. It is likely that there is a certain amount of truth in each of them, for different people, at different times, in different academic situations. This view has led to researchers trying to integrate some of the models, such as combining all goals and self-regulation (Urdan & Maehr, 1995; Wentzel, 1991) with personal investment across cultures (Maehr, 1984) and social and personal models (Weiner, 1994).

The ten models can be summarized under the following headings. One is the Arousal and Anxiety Model (Covington & Omelich, 1987; Naveh-Benjamin, 1991; Tobias, 1985). Two is a needs model (Darley, Glucksberg & Kinchla, 1988; Maslow, 1970). Three is an Achievement and Social Goal Model (Bandura, 1986; Maehr, 1984; McClelland, 1985; Urdan & Maehr, 1995; Wentzel, 1991). Four is a Behavioural Motivation Model involving rewards, reinforcement and intrinsic Motivation (Butler, 1988; Boggiano & Barrett, 1992; Lepper & Hodell, 1989; Heckhausen, 1991; Cameron & Pierce, 1994). Five is Attribution theory (Weiner, 1985; Maehr, 1989). Six is a Self-fulfilling Prophecy Model (Rosenthal, 1973: Good & Brophy, 1990). Seven is the Expectancy Value Model (Atkinson, 1964; Eccles et al., 1983). Eight is Self-regulated Learning Model (Corno, 1992; Reeve, 1996; Schunk, 1991; Wolters, 1998; Zimmerman, 1990; Zimmerman & Schunk, 1989). Nine is perceived self-efficacy model that relates personal beliefs to actions to achieve (Bandura, 1982; Schunk, 1989). Ten is a Personal Investment Model involving tasks, ego, social solidarity and extrinsic rewards (Maehr, 1984; Maehr & Braskamp, 1986).

In a recent evaluation of research on Motivation, Leo and Galloway (1996) called ‘for approaches to the study of Motivation which tap the phenomenology of the construct’ (p35) and they stated that ‘research in Motivation has yielded no consistent understanding about the nature or relevance of the construct’ (p.44)(by ‘construct’, they mean model of Motivation). Many researchers have not used a good multi-aspect model of Motivation and shown these aspects to be linked to behaviour.

For this paper, it was decided that the way ahead was to identify the main aspects and corresponding sub-aspects of Motivation from the ten models, operationally define the aspects by sub-aspects, and conceptually construct a scale by ordering the sub-aspects by ‘difficulty’ in Guttman-like patterns. Twelve main aspects of Motivation were found in the literature mentioned above and used in the present study to develop a new conceptual scale of Motivation. These aspects are interest, learning from others and responsibility for learning (as part of Desire to Learn), extrinsic, intrinsic and social rewards (as part of Personal Incentives), and standards, goals, tasks, effort, values and ability (as part of Striving for Excellence). So the present study attempts to develop a multi-aspect scale to measure Motivation and items were devised appropriate to each of the sub-aspects in Guttman-like patterns and with a direct link to learning behaviour.

Aims

There are three aims. One is to create an interval level, unidimensional scale of Motivation for university students, with attitude items linked to behaviour items, using a conceptual model of Motivation, based on Striving for Excellence (Standards, Goals, Tasks, Effort, Values and Ability), Desire to Learn (Interest, Learning from Others and Responsibility for Learning), and Rewards (Extrinsic, Intrinsic and Social). Two is to analyse its psychometric properties using the Extended Logistic Model of Rasch (Andrich, 1988a, 1988b; Rasch, 1980/1960) with the computer program Rasch Unidimensional Measurement Models (Andrich, Lyne, Sheridan, & Luo, 2000). Three is to investigate the structure and meaning of the Scale.

A New Scale of Motivation

A new scale of Motivation was conceptualised and created using ordered sub-groups of three main aspects of Motivation, linked with two response sets (What I aim for and What I actually do), and an ordered set of subject response categories (none or only one of my subjects, in some, though not most, of my subjects, in most, though not all of my, subjects, and in all or nearly all of my subjects). The model of Motivation was based on four simple ideas that when integrated would help explain Motivation as a complex variable.

The first involved the creation of a structure based on three main aspects of Motivation (Striving for Excellence, Desire to Learn and Personal Incentives), with each operationally defined by a number of sub-aspects. Striving for Excellence was defined by the sub-aspects Standards, Goals, Tasks, Effort, Ability and Values. Desire to Learn was defined by the sub-aspects Interest, Learning from Others, and Responsibility for Learning. Personal Incentives was defined by Extrinsic Rewards, Intrinsic Rewards and Social Rewards. The second involved creating stem-items in an ordered pattern by 'difficulty' within each sub-aspect. The structure of Motivation was then based on sub-sets of stem-items in patterns of ordered 'difficulty', each aligned from 'easy' to 'hard'. The third involved an ordered response set of categories for each of the stem-items. These are What I aim for (expected to be 'easy' on average) and What I actually do (expected to be 'harder' on average). What I aim for was expected to be stated by the students in terms of the number of subjects to which it applies. It was expected that this would reflect the students' needs, expectations, cognitions and desires, all internally and covertly contained within the students' minds, but now generally expressed in terms of what they aim for in their subjects. What I actually do was expected to be stated by the students in terms of the number of subjects to which it applies, too. It was expected that this would be decided by the students' personal beliefs, needs and cognitions, all now expressed as to what they actually do in their subjects. The fourth involved calibrating all the 'difficulties' of the items (from 'easy' to 'hard') onto the same scale as the measures of Motivation (from low to high), using a Rasch Measurement Model. The following material provides an example of the conceptual and model thinking involved with the construction of the scale for one of the sub-groups, Effort, under Striving for Excellence.

Expected Ordering by 'Difficulty' Pattern for the Effort Items

It was expected that most students would find it 'easy' to say that they aim to Make strong demands on themselves to achieve in academic work (item 15) in all their subjects. It was expected that there would be some variation in student responses around this. It was expected that most students would find it 'harder' to say that they aim to Make a strong effort to find the right answers, when they are given an academic task or assignment (item 17) in all their subjects and that there would be some variation around this. This is because item 17 involves 'a little bit more effort' conceptually than item 15. It was expected that most students would find it 'harder still' to say that they aim to Write and re-write their academic

assignments in order to achieve (item 19) and that there would be some variation around this. This is because item 19 involves 'a little bit more effort' conceptually than item 17. So it was expected that these three stem-items would form an ordered pattern of responses by 'difficulty', on average, from 'easy' to 'hard', when students reported that this is What I aim for. This is the vertical ordering of stem-items by 'difficulty' in the questionnaire set out in Appendix A.

Similarly, it was expected that this vertically ordered pattern of 'difficulties' for the students' self-views of What they aim for in relation to the three stem-items for Effort (as explained above) would be repeated for their self-views of their behaviour (What they actually do)(items 16, 18, 20). These patterns can be seen in the questionnaire (Appendix A) and in the sample below. That is, for the What I actually do self-view, the items would be ordered in 'difficulty' from 16 ('easiest') through 18 to 20 ('hardest').

Place Figure 1 about here

Expected Ordering by 'Difficulty' Patterns for the other Sub-groups

The expected order by 'difficulty' patterns for the other sub-groups follow a similar line to that reported above and are not reported here to avoid repetition and save space. A reader can easily work out the expected patterns of the other sub-groups from Appendix A. In Appendix A, the stem-items are placed in order from 'easy' to 'hard' in each sub-scale (see the logit values too) and the response categories (What I aim for and What I actually do) are also placed in order from 'easy' to 'hard'.

Measurement

There are many scales designed to assess Motivation (Ray, 1986, reported over 70 scales; see also Conoley & Impara, 1995, test numbers 226,244,245; Blankenship, 1987; Clarke, 1973; Fineman, 1977; Harper, 1975; Lian-Hwang Chiu, 1997; Piedmont, 1989; Thibert & Karsenti, 1996).

Many of these scales do not 'capture' a comprehensive view of Motivation and most involve a relatively simple range of aspects and items. These scales have all been analyzed with traditional measurement techniques and not with modern interval-level models, such as Rasch Measurement Models (Rasch 1980/1960). Lian-Hwang Chiu (1997) reported that reviews of the literature showed that the score reliability and validity of many of these scales varied from satisfactory to poor. The Twelfth, Eleventh, Tenth and Ninth Mental Measurements Yearbooks, as they relate to Motivation tests, support this conclusion. In addition, many of the scales are not based on a sufficiently detailed model of Motivation itself, nor linked to behaviour which is often part of their definition.

Problems with many Motivation assessments

Seven general aspects of many Motivation scales are called into question. First, most of the scales are not based on a 12 aspect model of Motivation (many scales had 4-6 aspects only). Two, most Motivation scales are not designed to measure the Motivation of a student who is highly motivated in one subject only and, at the same time, measure the Motivation of other students who are motivated to achieve in some or many subjects. Three, Likert (1932) response formats contain a discontinuity between the response categories of disagree and

agree. That is, the response measurement format is not ordered from low to high and those who are undecided, don't want to answer, are unclear or just neutral, will answer the middle (neutral) category. If a neutral category is not provided, they will be forced to answer either agree or disagree. This means there is a consequent interpretation problem. Four, researchers rarely test the linkage of their Motivation scales to behaviour, despite behaviour being linked to Motivation by definition. Hence, both What they aim for (Motivation) and What they actually do (behaviour), ought to be measured at the same time and calibrated on the same scale. Five, the items measuring Motivation to Achieve Academically are not always separated into their sub-scales on the questionnaires, so that it is not clear to the students what is being assessed. Six, positively and negatively worded items are often mixed to avoid the fixed response syndrome (a common procedure in traditional approaches). There is some evidence that this causes an interaction effect between items in modern measurement models (see Andrich & van Schoubroeck, 1989). Consequently, it is considered better to word all items in a positive sense when using modern measurement models, or to treat positively and negatively worded items as belonging to separate scales, unless they can be empirically shown to measure the same construct. Seven, the analysis of most Motivation scales has been performed with only traditional statistical programs and ordinal level scales. Modern measurement programs are now available to create interval level measures in which item difficulties and student Motivation measures can be calibrated on the same scale (Wright, 1985). They also test the conceptual structure of Motivation, including its dimensional nature (see Andrich, 1988a, 1988b; Andrich, Lyne, Sheridan & Luo, 1998; Rasch, 1960/1980; Waugh, 1998a,b). Rasch measurement model analysis has been shown as appropriate to use in measuring variables like Motivation (see Andrich, 1985, 1982; Waugh, 2001, 1999,1998a,b; Wright & Masters, 1982, 1981).

Changes made

The following changes were made to overcome the seven problems referred to above. The new stem-items for the scale were based on a model of three 1st order aspects defined by a number of 2nd order aspects. Forty-five items based on 12 main aspects of Motivation were devised to apply to students at university, with 45 corresponding behaviour items (see Appendix A). The items were ordered under their respective sub-group aspects that make it clear to the students what aspects are being assessed. They were all written in a positive sense, so as to be applicable to the new response format. The response format was changed in two ways. First, two columns were added for responses, one for What I aim for (measuring Motivation) and another for What I actually do (measuring behaviour). Second, the response categories were changed to an ordered format to provide an ordered measurement structure: in none or only one of my subjects, in some, though not most of my subjects, in most, though not all of my subjects, and in all or nearly all my subjects. This structure allowed for an enthusiast who is highly motivated in one or a few subjects, as well as for students who are motivated in some or many subjects. There are now 45 items relating to What I aim for and, in direct correspondence, 45 items relating to What I actually do (A sample is given below and the full instrument is given in Appendix A). The data were analyzed with a recent Rasch measurement model program (Andrich, Lyne, Sheridan, & Luo, 1998) to create a scale of Motivation and to test the conceptual and structural model of Motivation.

Please rate the 24 items according to the following response format and place a number corresponding to **What I aim for** and **What I actually do** on the appropriate line opposite each statement:

In all or nearly all my subjects put 3

In most, though not all, my subjects put 2

In some, though not most, of my subjects put 1

In none or only one of my subjects put 0

Example: If **you aim to set high standards** in academic work for all your subjects, put 3, and if **this only happens** in one subject, put 1.

Item 1. I set myself high standards in academic work. 3 1

Item no. Item wording What I What I

Aim for actually do

Sub-Scale: Striving for Excellence (20 items)

Standards

1/2 Do my best to reach the academic standards that I

set for myself.

3/4 Evaluate my performance against the academic

standards that I set myself.

Goals

5/6 Set myself realistic but challenging academic goals.

7/8 Try different strategies to achieve my academic goals

when I have difficulties.

Criteria for measurement

Seven measurement criteria have been set out by Wright and Masters (1981) for creating an interval level scale that measures a variable (see also Wright, 1985). Although these criteria are implicit in most ideas of scientific measurement, it is only with the advent of modern measurement models that they have been explicitly implemented in the social sciences. They are, 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 that is the same for all persons is required. Third, an evaluation of whether each person's responses form a valid response pattern is checked. Four, an estimation of each person's relative score (attitude or achievement) on the scale is created. 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. These criteria are used in the computer program RUMM to create the new scale.

Measurement Model

The Rasch method produces scale-free measures and sample-free item 'difficulties' (Andrich, 1988b; Wright & Masters, 1982). That is, mathematically, the differences between pairs of measures and pairs of item 'difficulties' are expected to be relatively sample independent in Rasch measurement. In contrast, Classical Test Theory, where the sums of scores on the items and the item 'difficulties' are not calibrated on the same scale, the totals are strictly sample dependent. Classical Test Theory cannot produce anything better than a ranking scale that will vary from sample to sample. The goal of a proper measurement scale for Motivation (a scale akin to a ruler) cannot be accomplished through Classical Test Theory.

Mathematically, in Rasch measurement, when all the items fit the model, there is a predominant single trait underlying all the items. This means that the measure estimated in Rasch modeling should be different from that calculated in Classical Test Theory. This could come about because the Rasch model will produce a different set of items contributing to the dominant trait than will Classical measurement. In recent times, DeMars (2001) and Fan (1998) have found that person ability estimates using Classical Test Theory gave comparable results to those using a Rasch Measurement Model. This implies that there is no need for a Rasch analysis, and that summing individual item scores in a variable is good enough. However, this is misleading. Fan (1998, p.368) also says that the results for the one-parameter Rasch model 'should be viewed with extreme caution' (p.368) because 30% of the items mis-fitted the one-parameter Rasch model. DeMars (2001) found similar results, with similar examples of mis-fitting items, and gives similar cautions. The central argument is that a Rasch measurement model (or something involving similar techniques) is needed to produce a proper scale in which both measures and item 'difficulties' are calibrated together, which Classical Test Theory does not do. Uses of Rasch analysis routinely do not count items that are not influenced by the unidimensional trait, thereby reducing the contribution of 'noise' to the measure. A counter claim is that content validity is decreased by deleting items and that Classical Test Theory then provides higher content validity by using more of the items. Rasch measurement requires the discrimination factor to be the same for all the items to form a proper scale. Many items are not included in Rasch analysis so that the discrimination condition can be fulfilled. Classical Test Theory allows the discrimination of the items to vary considerably. Factor analysis and inter-item correlations determine which items 'hang together' to produce scales in Classical Test Theory.

The zero point on the Rasch scale does not represent zero Motivation. It is an artificial point representing the mean of the item difficulties, calibrated by default to be zero, in Rasch measurement. It is possible to calibrate a true zero point, if it can be shown that an item represents zero Motivation. There is no true zero point in the present study. In Classical Test Theory there is no true zero point either, the difficulties of the items are not calibrated against the total scores, and the items are not conceptualised or calibrated to form a scale from 'easy' to 'hard'. In Rasch analysis, items are conceptualised and created initially in a scale from 'easy' to 'hard'. The data analysis tests this conceptualisation.

The RUMM program (1998) parameterises 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. When the thresholds are ordered

in line with the ordered response categories, the data fit the Rasch measurement model better. No check is normally made on this aspect in Classical measurement.

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 measures for each item and data are combined across all items to give an overall test-of-fit (see Andrich and van Schoubroeck, 1989, pp479-480 for the equations). This shows the collective agreement for all item locations across students of differing Motivation measures along the scale. No such check is normally done on this aspect in Classical measurement.

The item-student test-of-fit examines both the response patterns for students across items and for items across students. It examines the residual between the expected estimate and the actual values for each student-item summed over all items for each student and summed over all students for each item (see Styles and Andrich, 1993, p914 or Andrich and van Schoubroeck, 1989, p482 for the equations). The fit statistics approximate a distribution with a mean expected near zero and a standard deviation near one. Negative values indicate a response pattern that fits the model too closely (probably because response dependencies are present, see Andrich, 1985) and positive values indicate a poor fit to the model (probably because other measures ('noise') are present). Again, no checks are normally done on this aspect in Classical measurement.

Limitations

There are two main limitations to this study. They are [A] the acceptance or rejection of the Rasch measurement model and [B] the practice of measuring What I aim for and What I do simultaneously. With regard to the first limitation, one can use a model to fit the data (traditional approach) or model the data to fit strict measurement criteria (Rasch approach). The traditional approach has been to produce a complex model to fit the data using, for example, two person parameters (ability and guessing) and two item parameters (difficulty and discrimination) to form a model that would predict the data. The Rasch approach uses strict measurement criteria (described elsewhere in this paper) so that only items which fit the criteria can be ordered from 'easy' to 'hard' to form an interval level scale. These are the valid items measuring the variable. While there are disagreements about the two approaches (see Andrich, 1989; Divgi, 1986; Goldstein, 1980, 1979: Traub, 1983), the Rasch approach offers three advantages. First, it only uses items that fit the measurement criteria to form a valid measure of the variable. A check is made to see that persons respond to the valid items in a logical and consistent manner to form a scale and so 'noise' is considerably reduced. (In traditional approaches, groups of items as factors are identified, but no check is made that the item responses are answered in a logical and consistent pattern to form a scale). Second, the Rasch method creates an interval scale where items are ordered from 'easy' to 'hard' on the same scale as the Motivation measures. (This is not done in traditional measurement which only creates ordinal or ranking level measures). Third, it enables Motivation items and behaviour items to be calibrated on the same scale simultaneously and hence for a genuine link to be made between individual Motivation items and their corresponding learning behaviour. (Traditional techniques use correlations between overall Motivation and overall behaviour measures, with different Motivation and behaviour items. This typically produces low positive correlations in conflict with theory).

In the second limitation, the study assumes that students who are answering reliably about What they do, can at the same time remember What they aimed to do a week or so before. This may be questionable. On the one hand, there is evidence that students retrospective recollections may be biased by their implicit theories about personal change (Ross, 1989) and their memories may not be as reliable as they believe. Thus, it could be argued that it is better to measure What I aim forfirst and then measure What I do some weeks later. On the other hand, the questionnaire was trialled with 15 students who were questioned individually afterwards. The students said that they understood the questions, could easily remember What they aimed for and What they did and the questionnaire was to them valid and reliable.

Data collection

The sample consisted of 239 first year students selected through special entry programmes at an Australian University and was basically a convenience sample. There are 93 (39%) from SchoolSelect (where schools nominate students to attend university), 52 (22%) selected through UniStart Plus (fee-paying students who did not do sufficiently well in Tertiary Entrance Examinations for direct entry) and 94 (39%) selected through UniOps (where students were interviewed by university staff after application). All the students would have completed Year 12 at school the previous year and be turning 18 years old in the year of university entry. There were an equal mix of males and females. Their achievements would be at the lower end of university entrance in Western Australia. After ethics committee approval, the questionnaires were posted to students and a follow-up was conducted. For SchoolSelect, the response rate was 93/149 (62%), for UniStart Plus 52/60 (87%) and UniOps 94/280 (34%). Generally, the questionnaires took about 15 minutes to complete.

Results

The results are set out in one Figure, one Table and two Appendices. Figure 2 shows the graph of Motivation measures for the 239 students and the ‘difficulties’ of the 24 Motivation items (and their corresponding 24 behaviour items) on the same scale in logits. Table 1 gives a summary of the Index of Student Separation (proportion of observed variance considered true) and fit statistics for the 48 item scale. Appendix A shows the questionnaire items and the ‘difficulties’ of the 48 items. Appendix B shows, in probability order, the location on the continuum, fit to the measurement model and probability of fit to the model for the 48 items. Twenty-one of the original 45 stem-items were discarded (see Appendix A), because they had reversed thresholds indicating inconsistent category responses or the students could not agree on the ‘difficulty’ of the item on the scale (according to the fit statistics); that is, these items did not satisfy all of the criteria for measurement.

Place Table 1 and Figure 2 about here

Psychometric characteristics of the Motivation Scale

Twenty-four items relating to Motivation and 24 corresponding items relating to self-reported behaviour have a good fit to the measurement model, indicating a strong agreement between all 239 students to the different ‘difficulties’ of the items on the scale (see Table 1). 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 used. The Index

of Student Separability (akin to traditional reliability) for the 48 item scale is 0.928. This means that the proportion of observed variance considered true is 93%. The item-trait tests-of-fit indicate that the values of the item 'difficulties' are strongly consistent across the range of student measures ($p<0.0009$, see Table 1). The item-student tests-of-fit (see Table 1) indicate that there is good consistency of student and item response patterns. The items are appropriately targeted against the Motivation measures. That is, the range of item thresholds match the range of Motivation measures on the same scale (see Figure 2). The item threshold values range from –3.87 logits (Standard Error 0.2) to +3.55 logits (SE 0.2) and the Motivation measures of the students range from –1.9 logits (SE 0.2) to +6.8 logits (SE 0.2). There are only 10 students whose Motivation measures are more than +3.55 logits and hence not 'matched' against an item threshold on the scale. Taken together, these results indicate that a good measurement scale of Motivation has been created, that the data are reliable and consistent, that the errors are small in relation to the measures, and that the power of the tests-of-fit are excellent.

Discussion

Meaning of the Motivation Scale

The 24 stem-items that make up the variable Motivation are conceptualized from three 1st order orientations, operationally defined by a number of 2nd order orientations. The three 1st order orientations, Striving for Excellence, Desire to Learn and Personal Incentives are supported as contributing to the variable. The 2nd order orientations that involve standards, goals, tasks, effort and values are supported as contributing to Striving for Excellence (Ability is not supported). It is noted that items of the 2nd order aspect varied in difficulty from 'very easy' to 'hard', as expected. These are discussed later. The 2nd order orientations that involve Interest, Learning from Others and Responsibility for Learning are supported as contributing to Desire to Learn. The 2nd order orientations that involve Intrinsic Rewards and Social Rewards are supported as contributing to Personal Incentives (Extrinsic Rewards are not supported).

The reason that Extrinsic Rewards and Ability are not supported might be due to the wording of the items used in this study. There are very strong theoretical grounds for including Ability and Extrinsic Rewards as aspects of Motivation. It is suggested that a different wording be tried in a future trial of the scale.

The 24 stem-items that fitted the measurement model define the variable Motivation (see Appendix A). Each Motivation item is linked to a corresponding behaviour item such that it is 'easier' than the behaviour item on the scale. The items have good content validity and they are derived from a conceptual framework based on previous research. This, together with the data relating to reliability and fit to the measurement model (psychometric characteristics), is strong evidence for the construct validity of the variable. This can be held to mean that the students' responses to the 24 stem-items are related sufficiently well to represent the variable, Motivation.

The items of the scale are ordered from 'easy' to 'hard' (see Figure 1). Nearly all the students answered the 'easy' items positively for all their subjects. As the items become progressively 'harder' on the scale, the students need a higher Motivation to answer them positively for all subjects. Students with low measures of Motivation cannot answer the 'difficult' items positively for all subjects. The item 'difficulties' and the student measures are calibrated on the same scale at the interval level. Equal differences on the scale between measures of Motivation represent equal differences in item 'difficulty'. However, there is no true zero point of item 'difficulty' or Motivation. For the Motivation items (not the behaviour

items), the difficulties range from -1.79 (easiest item, 37 in Appendix A) to +0.69 (hardest item, 13 in Appendix A).

Conceptually valid, but non-fitting items

Twenty-one of the 45 stem-items that were considered to be conceptually valid did not fit the strict requirements of the Rasch measurement model and were discarded. However, the Rasch model does not tell the researcher how to fix the item to make it fit, if it doesn't fit. All it tells the researcher is whether the particular wording used for the item produces data that can be explained by a single predominant trait. The way forward is to reword the non-fitting items so that they fit the Guttman-type patterns in the model of Motivation conceptually and are consistent with the 24 stem-items already found to fit the model, and then test the model with another data set.

The relationship between items on the scale

Motivation items at the 'easiest' end of scale (for example 37,39,15,17 and 1, see Appendix A) are answered in agreement by nearly all the students. This means that students were likely to have found it 'easy' to say that they aimed to take personal responsibility for their academic learning and that they aimed to seek out information and take steps to master it, for all their subjects. They were likely to have found it 'easy' to say that they aimed to make strong demands on themselves to achieve in academic work and that they aimed to make a strong effort to find the right answers for academic assignments, for all their subjects.

Motivation items at the 'harder' end of the scale (for example 13,11,9, and 3, see Appendix A) are only answered in agreement by those students who have high measures of Motivation on the scale. It is most likely that students found it 'hard' to say that they aimed to seek some difficult academic tasks that they might be able to do and that they aimed to seek some difficult academic tasks in which they believe that they can succeed, for all their subjects. They found it 'hard' to say that they aimed to seek some average academic tasks in which they think they can succeed and that they aimed to evaluate their performance against the academic standards they set for themselves, for all their subjects.

The current analysis supports the conceptual design of Motivation as based on a multi-aspect model. That is, it supports the view that Motivation is based on an ordered line of three 1st order aspects (from Desire to Learn as the 'easiest', to Personal Incentives and Striving for Excellence as the 'hardest'). Each of these 1storder aspects is based on a number of 2nd order aspects, as previously stated. The 2nd order aspects for Desire to Learn vary from 'very easy' (responsibility for learning) to 'moderately easy' (learning from others) and 'moderately hard' (interest in academic work). The 2nd order aspects for Personal Incentives are 'moderately easy' (intrinsic and social rewards). The 2nd order aspects for Striving for Excellence vary from' very hard' (tasks) to 'very easy' (effort). In line with this, the analysis supports the view that the latent trait, Motivation, can be measured as a unidimensional variable based on the 1st order and 2nd order aspects.

Each of the Motivation items is linked to and falls at an 'easier' position on the scale than the corresponding behaviour item. Since the definition of Motivation involves internal processes that energize behaviour, this is evidence for the validity of the measure of Motivation and provides a good supporting test for this measure of Motivation. For example, with regard to Striving for Excellence, students found it 'easy' to say that they aimed to do their best to reach the academic standards that they set for themselves (item 1), but only 'moderately easy' to actually achieve it (item 2). They found it 'hard' to say that they aimed to seek some difficult tasks which they might be able to do (item 13) and 'very hard' to actually achieve this (item 14). They found it 'very easy' to say that they aimed to make strong demands on

themselves to achieve in academic work (item 15) and 'moderately hard' to say that they actually do so (item 16).

With regard to Desire to Learn, students found it 'very easy' to say that they aimed to show interest in a number of academic topics (item 23) and 'moderately easy' to actually do this (item 24). They found it 'easy' to say that they aimed to participate in class discussions to improve their understanding of academic matters (item 29) and 'moderately hard' to actually do this (item 30). They found it 'very easy' to aim to take personal responsibility for their academic learning (item 37) and 'moderately easy' to actually do this (item 38). They found it 'moderately hard' to aim to think about solving problems, with which others have difficulty, because they are interested (item 27) and 'very hard' to actually achieve this (item 28).

With regard to Personal Incentives, students found it 'moderately easy' to aim to achieve academically because of the challenges it brings (item 43) and 'moderately hard' to achieve this (item 44). They found it 'moderately hard' to aim to like the intellectual challenge of academic work (item 45) and 'much harder' to actually like the intellectual challenge of academic work (item 46). They found it 'moderately easy' to aim to like the social relationships involved in academic work (item 47) and 'much harder' to achieve this (item 48).

Summary

The Motivation scale brings together 10 of the 12 aspects that are often used in groups of 4 or 5 in other Motivation scales, as part of a multi-aspect model. This model is based on Striving for Excellence (Standards, Goals, Tasks, Effort, Values and Ability), Desire to Learn (Interest, Learning from Others and Responsibility for Learning), and Rewards (Extrinsic, Intrinsic and Social). Ability and Extrinsic Rewards do not fit the measurement model (the other 10 aspects do), probably because of the wording of the items. It is suggested that different item wordings for Ability and Extrinsic Rewards be tried in further research, as the conceptual model appears to be strong for the inclusion of these two aspects.

The measurement of Motivation is designed to include students who are highly motivated in one, two, three or more subjects. It places all 10 aspects on an interval level scale in which the student measures of Motivation and the item 'difficulties' are calibrated on the same scale. The items on the scale are ordered from 'easy' to 'hard' and the measures of Motivation are ordered from low to high. Individual Motivation items are shown to be related to corresponding self-reported behaviour items. The computer program RUMM (Andrich, Lyne, Sheridan & Luo, 1998) was very useful in creating and analyzing the scale.

References

Ames,C. & Ames,R. (1989)(Eds.). *Research on Motivation in education: Volume 3, Goals and cognitions*. San Diego,CA: Academic Press.

Ames,C & Ames,R. (1985)(Eds.). *Research on Motivation in education: Volume 2, The classroom milieu*. Orlando: Academic Press.

Ames,R. & Ames,C.(1984)(Eds.). *Research on Motivation in education: Volume 1, student Motivation*. New York: Academic Press.

Anderman, E.M. & Maehr, M.L.(1994). Motivation and schooling in the middle grades. *Review of*

Educational Research, **64** (2), 287-309.

Andrich,D.(1989). Distinctions between assumptions and requirements in measurement in the social

sciences. In J.A.Keats, R.Taft, R.A.Heath & S.Lovibond (Eds.), *Mathematical and Theoretical Systems,* pp. 7-16. Amsterdam (North-Holland):

Elsevier Science Publishers.

Andrich,D. (1988a). A General Form of Rasch's Extended Logistic Model for Partial Credit

Scoring. *Applied Measurement in Education,* **1**(4), 363-378.

Andrich,D.(1988b). *Rasch models for measurement*. Sage university paper on quantitative

applications in the social sciences, series number 07/068. Newbury Park, California: Sage

Publications.

Andrich,D.(1985). A latent trait model for items with response dependencies: Implications for test

construction and analysis. In S.E. Embretson (Ed.), *Test design: developments in psychology*

and psychometrics (pp. 245-275). Orlando: Academic Press.

Andrich,D.(1982). Using latent trait measurement to analyse attitudinal data: a synthesis of viewpoints. In D. Spearitt (Ed.), *The improvement of measurement in education and psychology,* (pp. 89-126). Melbourne: ACER.

Andrich,D., Lyne,A., Sheridan,B., & Luo,G. (1998). *RUMM: A windows-based item analysis*

program employing Rasch unidimensional measurement models. Perth: Murdoch University.

Andrich,D. & van Schoubroeck,L.(1989). The General Health Questionnaire: a psychometric

analysis using latent trait theory. *Psychological Medicine*, **19**, 469-485.

Atkinson,J.W.(1964). *An introduction to Motivation*. Princeton, NJ: Van Nostrand.

Bandura,A.(1986). *Social foundations of thought and action*. Englewood Cliffs, NJ: Prentice-Hall.

Blankenship,V.(1987). A computer-based measure of resultant achievement Motivation. *Journal of*

Personality and Social Psychology, **53**, 361-372.

Boggiano,A.K. & Barrett,M.(1992). Gender differences in depression in children as a function of

Motivational orientation. *Sex Roles*, **26**, 11-17.

Brophy,J.(1987). Synthesis of research on strategies for motivating students to learn. *Educational*

Leadership, **45**, 40-48.

Butler,R.(1988). Enhancing and undermining intrinsic Motivation: the effects of task involving and

ego-evolving evaluation on interest and performance. *British Journal of Educational*

Psychology, **58**, 1-14.

Cameron,J. & Pierce, W.D. (1994). Reinforcement, review and intrinsic Motivation: a meta-analysis.

Review of Educational Research, **64**(3), 363-423.

Clarke,D.E.(1973). Measures of achievement and affiliation Motivation. *Review of Educational*

Research, **43**, 41-51.

Conoley,J.C. & Impara,J.C.(1995)(Eds.). *The twelfth mental measurements yearbook.* Lincoln, NE:

Buros Institute of mental Measurements.

Corno,L.(1992). Encouraging students to take responsibility for learning and performance. *The*

Elementary School Journal**, 93**, 63-83.

Covington,M.V.(1992). *Making the grade: a self-worth perspective on Motivation and school*

reform. New York: Cambridge University press.

Covington, M.V. & Omelich, C.L. (1987). I knew it cold before the exam: a test of the anxiety

blockage hypothesis. *Journal of Educational Psychology*, **79**, 393-400.

Darley,J.M., Glucksberg,S. & Kinchla, R. (1988*). Psychology (4th ed.).* Englewood Cliffs,NJ:

Prentice-Hall.

DeMars, C. (2001). Group differences based on IRT scores: Does the model matter? *Educational*

And Psychological Measurement, **61** (1), 60-70.

Divgi,D.R.(1986). Does the Rasch model really work for multiple choice items? Not if you look

closely. *Journal of Educational Measurement*, **23** (4), 283-298.

Donohue,T.L. & Wong,E.H. (1997). Achievement Motivation and college satisfaction in traditional

and nontraditional students, *Education,* **118**, 237-143.

Duke,M.P. & Nowicki,S. (1974). Locus of control and achievement: the confirmation of a theoretical

expectation. *The Journal of Psychology,* **87**, 263-267.

Eccles,J., Adler,T., Fulterman,R., Goff,S., Kaczala,C., Meece,J. & Midgley,C. (1983). Expectancies,

values and academic behaviours. In J.T. Spence (Ed.), *Achievement and achievement*

Motivation (pp. 75-146). San Francisco,CA: W.H.Freeman.

Fan, X. (1998). Item response theory and classical test theory: Am empirical comparison of their

item/person statistics. *Educational and Psychological Measurement*, **58**, 357-381

Fineman,S.(1977). The achievement motive construct and its measurement: where are we now?

British Journal of Psychology, **68**, 1-22.

Friis,R.H. & Knox,A.B. (1972). A validity study of scales to measure need achievement, need

affiliation, impulsiveness and intellectuality. *Educational and Psychological Measurement,*

32,147-154.

Goldstein,H.(1980). Dimensionality, bias, independence and measurement scale problems in latent

trait test score models. *British Journal of mathematical and Statistical Psychology*,

33, 234-246.

Goldstein,H.(1979). Consequences of using the Rasch model for educatioanl assessment. *British*

Journal of EducationalResearch, **5**, 211-220.

Good,T.L. Brophy,J.E.(1990). *Educational psychology: a realistic approach (4th ed.)*. White Plains,

NY: Longman.

Graham, S. (1994). Motivation in African Americans. *Review of Educational Research,* **64**(1), 55-

117.

Harper,E.B.W.(1975). The validity of some alternative measures of achievement Motivation.

Educational and Psychological Measurement, **40**, 531-536.

Heckhausen,H. (1991). *Motivation and action*. Berlin: Springer-Verlag (translation by Peter K.

Leppmann).

Helmreich,R.L. & Spence,J.T. (1978). The Work and Family Orientation Questionnaire: An

objective instrument to assess components of achievement Motivation and attitudes towards

family and career. *JSAS Catalogue of Selected Documents inPsychology,* **8**, 35.

Jacobs, P.A. & Newstead, S.E. (2000). The nature and development of student motivation. *British*

Journal of Educational Psychology, **70** (2), 243-253.

Leo,E.L. & Galloway,D.(1996). Evaluating research on Motivation: generating more heat than light.

Evaluation and Research in Education, **10**(1), 35-47.

Lepper,M.R. & Hodell,M. (1989). Intrinsic Motivation in the classroom. *Research on Motivation in*

the Classroom, **3**, 73-105.

Lian-Hwang Chiu, (1997). Development and validation of the school achievement Motivation rating

scale. *Educational and Psychological Measurement,* **57**, 292-305.

Likert,R.(1932). A technique for the measurement of attitudes. *Archives of Psychology*,**140**. Entire issue.

Maehr,M.L.(1989). Thoughts about Motivation. In C.Ames & R.Ames (Eds.), *Research on Motivation in education: Volume 3, Goals and cognitions*. Orlando: Academic Press.

Maehr, M.L. (1984). Meaning and Motivation: toward a theory of personal investment. In R.E.Ames and C.Ames (Eds.). *Research on Motivation in education* (Vol.1, pp. 115-144). New York: Academic Press.

Maehr,M.L. & Ames,C. (1989). *Advances in Motivation and achievement: volume 6 motivating enhancing environments*. Greenwich, C.T.: JAI Press.

Maehr, M.L. & Braskamp, L.A.(1986). *The Motivation factor: a theory of personal investment*. Lexington, Mass.: Lexington Press

Maehr,M.L. & Pintrich,P.R. (1997)(Eds.).*Advances in Motivation and achievement (Volume 10)*. London: JAI Press.

Maehr,M.L. & Pintrich,P.R. (1995)(Eds.).*Advances in Motivation and achievement (Volume 9): Culture, Motivation and achievement*. Middlesex,UK: JAI Press.

Maehr,M.L. & Pintrich,P.R. (1993)(Eds.). *Advances in Motivation and achievement (Volume 8):Motivation and adolescent development*. Middlesex,UK: JAI Press.

Maslow,A.H.(1970). *Motivation and personality (2nd ed.).* New York: Harper & Row.

McClelland,D.C.(1985). How motives, skills and values determine what people do. *American Psychologist,* **40**, 812-825.

Naveh-Benjamin,M.(1991). A comparison of training programs intended for different types of test-anxious students: further support for an information processing model. *Journal of Educational Psychology,* **83**, 134-139.

Peidmont,R.L.(1989). The Life Activities Achievement Scale: An act-frequency approach to the measurement of Motivation. *Educational and Psychology Measurement*,**49**, 863-874.

Rasch,G.(1980/1960). *Probabilistic models for intelligence and attainment tests (expanded edition)*.

Chicago: The University of Chicago Press (original work published in 1960).

Ray,J.J.(1986). Measuring achievement Motivation by self-reports. *Psychological Reports,* **58**, 525-

526.

Reeve, Johnmarshall (1996). *Motivating others: nurturing innerMotivational resources*. Needham

Heights, MA: Allyn & Bacon

Rosenthal,R. (1973). The Pygmalion effect lives. *Psychology Today*, September 56-73.

Ross,M.(1989). Relation of implicit theories to the constructionof personal histories. *Psychological*

Review, **96**, 341-357.

Schunk, D.H.(1991). Self-efficacy and academic Motivation. *Educational Psychologist,* **26**, 207-231.

Schunk, D.H. (1990) (Sp. Ed.). Motivation and Efficacy. *Journal of Educational Psychology*, **82**,1.

Styles, I. & Andrich, D. (1993). Linking the standard and advanced forms of the Raven's Progressive

matrices in both the pencil-and-paper and computer-adaptive testing formats. *Educational*

and Psychological Measurement, **53** (4), 905-925.

Thibert,G. & Karsenti,T.P. (1996). *Motivation profile of adolescent boys and girls: gender*

differences throughout schooling. Paper presented at the Annual Conference of the American

Educational Research Association in San Francisco, CA, April 8-12, 1996.

Tobias,S.(1985). Test anxiety: interference, defective skills and cognitive capacity. *Educational*

Psychologist, **20**, 135-142.

Traub,R.E.(1983). A priori considerations in choosing an item response. In R.K.Hambleton (Ed.), pp

57-70, *Applications of item response theory*.

Urdan,T.C. & Maehr,M.L.(1995). Beyond a two-goal theory of Motivation and achievement: a case

for social goals. *Review of Educational Research,* **65**(3), 213-243.

Waugh, R.F. (2001). Measuring ideal and real self-concept on the same scale, based on a multi-

faceted, hierarchical model of self-concept. *Educational and Psychological Measurement*, **61**

(1), 85-101.

Waugh, R.F. (1999). Approaches to studying for students in higher education: A Rasch measurement

model analysis. *British Journal of Educational Psychology*, **69** (1), 63-80.

Waugh,R.F. (1998a). The Course Experience Questionnaire: A Rasch measurement model analysis.

Higher Education Research and Development, **17** (1), 45-64

Waugh,R.(1998b). *A Rasch measurement model analysis of an Approaches to Studying Inventory for*

students in higher education. Paper presented at the Latent Trait Theory Conference: Rasch

measurement, held at the University of Western Australia from January 22-24, 1998.

Weiner,B.(1994). Integrating social and personal theories of achievement striving. *Review of Educational Research,* **64** (4), 557-573.

Weiner,B.(1990). History of Motivational research in education. *Journal of Educational Psychology*, **82**, 4, 616-622.

Weiner,B.(1985). An attributional theory of Motivation and emotion. *Psychological Review*, **92**, 548-573.

Wentzel, K. (1991). Social and academic goals at school: Motivation and achievement in context. In P.Pintrich & M.L.Maehr (Eds.). *Advances in Motivation and achievement (Vol.7)*, Goals and self-regulatory processes (pp 185-212). Greenwich,CT: JAI Press

Wolters, C.A. (1998). Self-Regulated learning and college students' regulation of Motivation. *Journal of Educational Psychology*, **90** (2), 224-235

Wright,B.D.(1985). Additivity in psychological measurement. In E.E. Roskam(Ed.), *Measurement and Personality Assessment,* pp101-112.Amsterdam, North-Holland: Elsevier Science

Wright,B. & Masters G.(1982). *Rating scale analysis: Rasch measurement.* Chicago: MESA press.

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

Zimmerman,B.J. (1990). Self-regulated learning and academic achievement: an overview. *Educational Psychologist,* **25**, 3-17

Zimmerman,B.J.,Bandura,A. & Martinez-Pons,M. (1992). Self-Motivation for academic attainment: the role of self-efficacy beliefs and personal goal setting. *American Educational Research Journal*, **29**, 663-667.

Zimmerman,B.J .& Schunk,D.H.(1989)(Eds.). *Self-Regulated learning and academic achievement:*

theory, research and practice. New York: Springler-Verlag

Appendix A: **QUESTIONNAIRE: MOTIVATION TO ACHIEVE ACADEMICALLY**

Please rate the 48 items according to the following response format and place a number corresponding to **What I aim for** and **What I actually do** on the appropriate line opposite each statement:

In all or nearly all my subjects put 3

In most, though not all, my subjects put 2

In some, though not most, of my subjects put 1

In none or only one of my subjects put 0

Example: If **you aim to set high standards** in academic work for all your subjects, put 3, and if **this only happens** in some subjects, put 1.

Item 1. I set myself high standards in academic work. 3 1

Item no. Item wording What I What I

Aim for actually do

Sub-Scale: Striving for Excellence (11 stem-items fit the model)

Standards

1/2 Do my best to reach the academic standards that I

set for myself. -0.97 -0.55

3/4 Evaluate my performance against the academic

standards that I set myself. +0.14 +0.45

Set myself the highest standards in academic work

Which I believe I can achieve. Did not fit the model

Goals

5/6 Try different strategies to achieve my academic goals

when I have difficulties. -0.07 +0.69

7/8 Set myself realistic but challenging academic goals. -0.28 -0.11

Set the highest academic goals which I can achieve. Did not fit the model

When I have difficulties in reaching my goals, I make

a renewed effort to ensure I achieve my goals. Did not fit the model

Tasks

9/10 Seek some average academic tasks in which I

think I can succeed. +0.38 +0.52

11/12 Seek some difficult academic tasks in which

I believe I can succeed. +0.45 +1.26

13/14 Seek some difficult academic tasks which I

might be able to do. +0.67 +1.35

Seek some easy academic tasks in which I am

Strongly likely to succeed. Did not fit the model

Seek some easy academic tasks which I might

be able to do. Did not fit the model

Effort

15/16 Make strong demands on myself to achieve in

academic work. -1.14 +0.30

17/18 When I am given an academic task or assignment,

I make a strong effort to find the right answers. -1.05 -0.66

19/20 Write and re-write my academic assignments

in order to achieve. -0.09 +0.75

Prepare myself to achieve as high as I can

in my academic assignments. Did not fit the model

Make a strong effort to achieve as high as I can

in academic work. Did not fit the model

Values

21/22 When I have conflicts about time to be spent on

achieving, I re-think my values (social, parental,

dates versus achievement). -0.10 +0.71

Value achievement in academic work. Did not fit the model

Ability

Have confidence in my academic ability to achieve the

best that is possible with my ability. Did not fit the model

Have positive feedback from my teachers on my ability

in academic work. Did not fit the model

Have positive feedback from at least one peer

Friend on my ability in academic work. Did not fit the model

Have positive feedback from at least one parent (or

guardian) on my ability in academic work. Did not fit the model

Sub-Scale: Desire to learn (9 stem-items fit the model)

Interest

23/24 Show interest in a number of academic topics. -0.94 -0.38

25/26 Read widely on a number of academic topics. +0.08 +1.08

27/28 Think about solving problems, with which others have

difficulty, because I'm interested. +0.12 +0.85

Display curiosity about the world and 'how it works'. Did not fit the model

Behave conscientiously in my academic work. Did not fit the model

Learning from others

29/30 Participate in class discussions to improve

my understanding in academic matters. -0.71 +0.61

31/32 Ask questions of others to improve my

understanding in academic matters. -0.69 +0.36

33/34 Learn from others with more knowledge than I have. -0.49 +0.00

35/36 Aim to learn from an expert in at least one

academic area. -0.21 +0.36

Try to pay attention to my teachers in order to

learn as much as I can. Did not fit the model

Responsibility for learning

37/38 Take personal responsibility for my academic

learning. -1.68 -0.35

39/40 Plan to seek out information when necessary and

take steps to master it. -1.47 -0.33

Sub-Scale: Personal Incentives (4 stem-items fit the model)

Extrinsic Rewards

Try to achieve academically because I like the

rewards it brings to me. Did not fit the model

Try to achieve academically because I like the

status it brings to me. Did not fit the model

Try to achieve academically because I like the

Competition with others that it brings. Did not fit the model

Intrinsic Rewards

41/42 Like the interaction with peers in solving

problems in academic work. -0.18 +0.35

43/44 Try to achieve academically because I like the

challenges it brings. -0.11 +0.63

45/46 Like the intellectual challenge of academic work. +0.08 +0.64

Like the curiosity of academic work. Did not fit the model

Social Rewards

47/48 Like the social relationships involved in academic

work. -0.34 +0.08

Have fun with others while involved in academic work. Did not fit the model

Bring honor to my family by succeeding in academic Did not fit the model

work.

Notes on Appendix A

1. Item ‘difficulties’ are in logits (log odds of answering positively)

2. The lower the item ‘difficulty’, the ‘easier’ the item

3. The higher the item ‘difficulty’, the ‘harder’ the item

4. Items at the ‘easiest’ end of the scale are answered positively by almost all the students. Students need a high Motivation to answer the ‘hardest’ items positively.

5. The What I aim for items are ‘easier’ than the What I actually do items.

6. The standard errors vary from about 0.09 to 0.20 (see Appendix B)

Appendix B: **Location (item 'difficulty') and Probability of item fit for** Motivation **(24 items) and corresponding behaviour (24 items)**

Item Location SE Residual ChiSq Probability

I0029 -0.715 0.12 0.096 0.130 0.988

I0038 -0.346 0.11 0.796 0.178 0.980

I0020 +0.754 0.09 0.398 0.338 0.951

I0041 -0.180 0.11 -0.251 0.710 0.867

I0008 -0.115 0.10 0.407 0.772 0.852

I0042 +0.353 0.09 0.242 0.779 0.850

I0034 +0.003 0.11 0.422 0.900 0.820

I0032 +0.356 0.09 0.043 1.162 0.755

I0036 +0.364 0.09 -0.277 1.429 0.689

I0031 -0.687 0.14 -0.786 1.445 0.686

I0002 -0.545 0.11 -0.173 1.849 0.592

I0017 -1.052 0.15 -1.064 2.166 0.525

I0025 +0.079 0.10 0.512 2.367 0.484

I0022 +0.707 0.09 -0.551 2.384 0.481

I0019 -0.092 0.10 -0.069 2.578 0.445

I0007 -0.283 0.11 0.483 2.658 0.430

I0006 +0.690 0.09 -0.498 2.762 0.412

I0023 -0.935 0.13 -1.168 2.863 0.395

I0035 -0.210 0.11 -0.460 2.987 0.375

I0026 +1.080 0.10 0.136 3.030 0.368

I0018 -0.659 0.11 -0.341 3.049 0.365

I0011 +0.448 0.09 1.305 3.086 0.359

I0003 +0.142 0.10 2.486 3.101 0.357

I0033 -0.490 0.13 0.231 3.138 0.352

I0040 -0.334 0.10 -0.910 3.348 0.321

I0039 -1.468 0.15 -1.558 3.890 0.251

I0037 -1.675 0.20 -1.424 3.992 0.240

I0012 +1.263 0.10 0.643 4.247 0.212

I0001 -0.971 0.15 0.635 4.423 0.195

I0024 -0.378 0.12 -0.030 4.636 0.176

I0004 +0.449 0.09 1.159 4.638 0.176

I0048 +0.076 0.10 1.053 5.690 0.101

I0028 +0.848 0.09 0.292 5.735 0.098

I0016 +0.298 0.10 -1.294 5.925 0.088

I0021 -0.098 0.10 1.084 5.926 0.088

I0047 -0.343 0.11 0.903 5.977 0.085

I0014 +1.347 0.10 0.609 6.004 0.084

I0045 +0.079 0.10 -1.557 6.756 0.052

I0043 -0.110 0.11 -1.286 6.772 0.051

I0005 -0.074 0.11 -1.437 6.873 0.048

I0027 +0.120 0.10 1.302 7.256 0.035

I0044 +0.627 0.09 -1.464 8.621 0.005

I0015 -1.140 0.13 -1.554 8.738 0.003

I0030 +0.608 0.09 2.476 8.820 0.002

I0013 +0.672 0.09 1.488 9.208 0.000

I0046 +0.636 0.10 -1.711 11.530 0.000

I0010 +0.520 0.10 2.140 14.136 0.000

I0009 +0.379 0.09 2.780 18.253 0.000

==========

Notes on Appendix B

1. Location is the item ‘difficulty’ in logits

2. SE is the standard error in logits

3. Residual is the difference between the observed and expected responses

4. Probability is the chi-square fit to the measurement model. Chi-square is sensitive to sample size and is not to be taken too strictly for large samples, like this study.

What I What I aim for actually do

Effort

15/16 Make strong demands on myself to achieve in

academic work. -1.14 +0.30

17/18 When I am given an academic task or assignment,

I make a strong effort to find the right answers. -1.05 -0.66

19/20 Write and re-write my academic assignments

in order to achieve. -0.09 +0.75

Figure 1. Structure of Effort items by ‘difficulty’

LOCATION PERSONS ITEMS [uncentralised thresholds]

7.0 |

X |

|

|

|

6.0 |

|

|

|

|

5.0 |

X |

|

|

|

4.0 X |

X |

X |

X | I0014.3

XXXX | I0026.3 I0012.3

3.0 XX |

XX |

XXX | I0046.3

XXXXXXX | I0024.3 I0002.3

XXXXX | I0006.3 I0044.3 I0016.3

2.0 XXXXXXXXXX | I0011.3 I0022.3 I0004.3 I0010.3

XXXXXXXXXXXXXX | I0020.3 I0040.3 I0013.3 I0028.3

XXXXXXXXX | I0042.3 I0032.3 I0030.3 I0008.3

XXXXXXXXXXX | I0012.2 I0045.3 I0036.3

XXXXXXXXXXXXXXXXX | I0025.3 I0034.3

1. XXXXXXXX | I0004.3 I0026.2 I0030.2 I0020.2 I0014.2 I0018.3

   I0009.3 I0048.3

XXXXXXXX | I0046.2 I0044.2 I0004.2 I0021.3 I0043.3 I0005.3

XXXX | I0016.2 I0028.2 I0023.3 I0038.3 I0022.2 I0027.3

XXXX | I0047.3 I0009.2 I0036.2 I0011.2 I0006.2 I0007.3

I0019.3 I0041.3 I0015.3

XXX | I0008.2 I0029.3 I0042.2 I0040.2 I0048.2 I0010.2

1. XXX | I0027.2 I0035.2 I0025.2 I0019.2 I0035.3 I0013.2

   I0032.2

X | I0007.2 I0002.2 I0041.2 I0013.1 I0033.3 I0003.2

I0028.1 I0039.3 I0031.3

X | I0005.2 I0043.2 I0045.2 I0021.2 I0017.3 I0047.2

X | I0034.2 I0018.2 I0024.2 I0027.1 I0038.2 I0029.2

I0003.1 I0009.1 I0001.3

X | I0034.1 I0005.1 I0032.1 I0033.1 I0022.1 I0006.1

I0033.2 I0020.1 I0014.1

-1.0 | I0015.2 I0010.1 I0036.1 I0012.1 I0031.2 I0041.1

I0019.1 I0043.1 I0021.1 I0001.2 I0035.1
I0030.1

I0042.1

| I0025.1 I0026.1 I0031.1 I0037.3 I0045.1

| I0037.2 I0017.2 I0048.1 I0047.1 I0011.1 I0038.1

I0007.1 I0044.1

| I0001.1 I0017.1 I0039.2

X | I0023.1 I0023.2 I0004.1 I0046.1

-2.0 | I0029.1

| I0016.1

| I0008.1

| I0018.1

| I0037.1

-3.0 | I0039.1

| I0040.1 I0024.1 I0015.1

|

|

|

-4.0 | I0002.1

Figure 2: Scale of Motivation to Achieve Academically

Notes on Figure 2

1. The scale is in logits, the log odds of answering positively.
2. Measures of Motivation are calibrated on the same scale as the item difficulties.
3. The items are towards the 'easy' end of the scale compared to the Motivation measures. This means the 'test' was a little 'easy' for these students.
4. Items at the easy end of the scale are answered positively by most students. As the

items become 'harder', students need a higher Motivation to answer the items

positively.

5. Each X represents 2 students.

6. The 'difficulties' are the thresholds. For example I0014.3 ('hardest' item) is the

threshold between the 3rd and 4th category responses for item 14. I0002.1 is the

threshold between the 1st and 2nd category responses for item 2.

Table I

Summary data of the reliabilities and fit statistics to the model for the 48 item Motivation scale (N = 239)

Non-fitting items none

Items with disordered thresholds none

Items with residuals -2>x>+2 48

Index of Student Separability 0.928

Item mean (SD) 0.000 (1.149)

Student mean (SD) 1.652 (0.946)

Item-trait interaction (chi-square) 217 ($p<0.0009$)

Item fit statistic mean +0.089

Sd +1.149

Student fit statistic mean -0.246

Sd +1.558

Power of test-of-fit excellent

Notes on Table I

1. The Index of Student Separation is the proportion of observed variance that is considered true (93%) and is high.
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 between students of differing Motivation for all item difficulties.