Teacher Perceptions of Learned Helplessness:
A Rasch Analysis of the Student Behaviour Checklist

Shirley M. Yates
Tilahun Mengesha Afrassa
The Flinders University of South Australia

In a longitudinal study of third grade students Fincham, Hokoda and Sanders (1989) reported that teacher perceptions of learned helplessness had a strong and consistent relation to concurrent achievement and to mathematics achievement test scores two years later. Helplessness was measured with the Student Behavior Checklist, in which teachers rated the student behaviour on a five point scale on 12 items measuring learned helplessness and 12 items measuring mastery. While the two subscales were highly correlated ($r = -0.81$), Fincham et al (1989) reported that the checklist was not a robust psychometric instrument. Under an assumption of the importance of teacher detection of learned helplessness and the relationship between this helplessness and school achievement, a study was conducted in 1994 with 58 teachers in 31 schools who rated 258 primary and lower secondary students in South Australia. This project aimed to assess the psychometric properties of the instrument and to determine if teachers could in fact detect learned helplessness through the use of the Student Behavior Checklist. The resultant data were analysed with confirmatory factor analyses and with the Rasch procedure. The results indicated that, while teachers could not judge learned helplessness as such, a unidimensional scale of academic behaviour based on a modified Student Behavior Checklist did allow for useful discriminations to be made amongst students.


Introduction

The present study is part of a continuing investigation into motivational variables implicated in primary school students' mathematics achievement. Data from this project have been presented at the Australian Association for Research in Education conferences in
Fremantle (1993) and Newcastle (1994), and have been published in Educational Psychology (Yates, Yates & Lippett, 1995) and the Social Psychology of Education (Yates, Keeves & Afrassa, in press, 1996). Additional data are being collected currently in order to conduct longitudinal analyses of students over a three year period.

Amongst the data collected were teacher ratings on student classroom characteristics. These data included measures of academic helplessness, as perceived by teachers, with a view to entering this variable into additional analyses.

Although the concept of learned helplessness now has a long history in psychology, there appears to be no recognised measure of this trait in terms of teachers' perceptions and judgements. Helplessness is often defined by the use of student self report indices such as the Intellectual Achievement Responsibility Scale (IAR; Crandall, Katovsky & Crandall, 1965), by various attributional-type line scales or by the Children is Attributional Style Questionnaire (Seligman, Peterson, Kaslow, Tannenbaum, Alloy, & Abramson, 1984). In the current project the use of a teacher-rating instrument, one that emerged from the work of Fincham, Hokoda and Sanders (1989) was investigated. The internal psychometric properties of this instrument were examined on the assumption that the development of a short and acceptable rating scale would be an important step for both research and clinical practice in this area.

Learned Helplessness

Helplessness is described by Peterson, Maier and Seligman (1992) in terms of three criteria: (a) loss of motivation, (b) changes in cognition and emotion, and (c) a reduction in behavioural agency (such as passivity). Amongst the changes in cognition is the perception of non-contingency; that is, the belief that important outcomes are uncontrollable. Helplessness in children has been predominantly measured by self report using paper and pencil. Thus the available research has largely used students in the fifth grade or higher.

In classroom contexts it is likely that helplessness is observed through the way students respond to situations of actual or conceivable failure. It may be thus assumed that teachers are in a position to assess at least some of the recognised dimensions of helplessness as they surface in classroom life.

The Student Behavior Checklist

In developing the Student Behavior Checklist that was used in this study, Fincham et al. (1989) generated items that reflected the range of behaviours associated with learned helplessness and mastery...
orientation in the research literature. Thus, by their very nature the items reflect student characteristics that are directly observable by teachers, rather than being inferred from an internal state as measured in student self reports. Fincham et al (1989) reported that although the learned helplessness and mastery orientation subscales were highly correlated ($r = 0.81$) the psychometric robustness of the instrument had yet to be established. Furthermore they raised the issue as to whether the scales specifically measured learned helplessness and mastery orientation or whether they reflected academic competence. Lastly, they felt that as the scale was strongly related to concurrent and future achievement scores in their own study and that of Nolen-Hoeksema, Girgus and Seligman (1986), perhaps a shorter version of the scale may "provide a cost-effective measure of helplessness" (Fincham et al, 1989, p 143).

Teacher Judgement

1Classroom behaviour

In a critical review of teacher-administered rating scales of the classroom behaviour of children Spivack and Swift (1973) noted the importance of ascertaining student behavioural adjustment in the classroom not only from a behavioural management point of view but also because it reflected "the extent to which the child may be benefitting from participation in the educational enterprise itself" (Spivack & Swift, 1973, p55). In reviewing the literature of the time they found 19 studies in which teachers had rated overt behaviours, and in most of these there was both a paucity of classroom behaviours covered and a marked lack of psychometric rigour in the scales themselves. With respect to teachers as judges, they reported that teacher ratings discriminated between a variety of criteria, had some stability over time, and that teachers' ratings of girls' overt behaviour were more consistent with their actual performance than was the case for boys. In stressing that need for more research in this area they stated that there is reason to believe that certain overt behaviours of the child in the classroom bear an intimate relationship to his adaptive capacity and consequent achievement in the setting, and that these relate more highly to his academic success than do general dimensions of adjustment or personality functioning (Spivak & Swift, 1973, p87). It was considered that the study of overt student behaviour by teachers supplied a new dimension to the understanding of classroom behaviour and school achievement.

2Academic performance

Hoge and Coladarci (1989) located 16 studies in which teachers' judgements of their students' academic performance were compared against actual scores on objective test measures. Across the studies
the median correlation was 0.66 suggesting a strong correspondence between teacher judgements and student achievement. The data from several studies suggested that teachers achieved a 'hit-rate' of around 70 per cent accuracy when asked to assess whether individual students are able to succeed on specific test items.

When the judgements of teachers were compared Hoge and Coladarci noted that a number of studies indicated large variations amongst individual teachers. For example, Hoge and Butcher (1984) reported correlations of between 0.4 and 0.87 across a sample of 12 teachers. Moreover, they reported that the accuracy of teacher judgements appeared to be relatively higher in the case of judgements made on average to above average ability children.

The overall conclusion of the Hoge and Coladarci review was that, with regard to the achievement domain, teacher judgements did concur with more objective measures. However some teachers tended to be more accurate than others and there was a tendency for teachers to err in over-estimating the capabilities of low-achieving students.

3 Teacher grading

In a review of 19 studies of teacher grading over the last ten years Brookhart (1994) also noted variability in teacher practices. Different teachers not only perceived the meaning and purposes of grades differently, but considered achievement and nonachievement factors differently (Brookhart, 1993; Frary, Cross & Weber, 1993; Nava & Lloyd, 1992; Pilcher-Carlton & Oosterhof, 1993). With respect to achievement and nonachievement factors, Brookhart noted the confounding effect of effort and achievement on teachers' grading. When grading student work, teachers see effort as a separate issue from considering a student's gender or personality (Frary et al, 1993; Griswold & Griswold, 1992; Gullickson, 1985; Nava and Lloyd, 1992; Pilcher-Carlton & Oosterhof, 1993; Stiggins, Frisbe & Griswold, 1989; Wood, Bennett, Wood & Bennett, 1990). These comments are important as the characteristics of learned helplessness include passivity, loss of motivation and lack of effort, behaviours which in turn impact on academic achievement. If students do not participate in the activities and lessons provided by the teachers, then their achievement is jeopardised (Brookhart, 1994). She also noted a difference between primary and secondary teachers, with the former relying more on observation and informal evidence and the latter more on written evidence when grading.

In a study of a random sample of secondary school teachers of academic subjects, Frary et al (1993) studied their grading practices and found tremendous variability. Though the use of cluster analysis they identified six groups of teachers, whose patterns of responding
indicated underlying differences in their approaches to grading. The smallest group used a norm referenced approach to grading, with the five other groups being respectively classified as soft-hearted, strict, arbitrary, uncertain and inconsistent. These clusters and their descriptions are described by Brookhart as being a major contribution to research in this field.

The Present Study

The extent to which teachers' judgements of student emotional and motivational traits reflect the high level of accuracy that is apparent within the achievement domain is of course open to question. Knowledge of this area has been hampered by lack of suitable measurement instruments. Thus the present study sought to investigate properties of the Fincham, Hokoda & Sanders (1989) Student Behavior Checklist and the extent to which this scale could be used by teachers to detect learned helplessness. This scale was chosen for investigation because of its importance in the literature in investigations of student achievement and explanatory style (Nolen-Hoeksema, Girkus & Seligman, 1986).

METHOD
Subjects

Fifty-eight teachers in 31 schools in an Australian city rated 258 students from Years 4 - 8 with the Student Behavior Checklist. The ratings were made in November 1994.

Materials

The Student Behaviour Checklist developed by Fincham et al (1989) is comprised of 24 items, 12 of which were selected from the research literature to measure the construct of learned helplessness, while the other 12 were designed to measure mastery orientation. An example of an item measuring learned helplessness is "Prefers to do easy problems rather than hard". An example of an item measuring mastery orientation is "Tries to finish assignments even when they are difficult". The teachers were asked to rate student behaviour over the past two to three months on a five point scale ranging from 1 (not true) to 5 (very true).

Procedure

In Term 1, 1993, 293 students in two primary schools took part in a study of the relationship between mathematics achievement and explanatory style. In Term 4, 1994, 258 of these students were traced to 31 different schools. Each school was contacted initially by telephone and the teachers invited to complete the Student Behavior Checklist which was then forwarded to them. Fifty-eight teachers in
these 31 schools completed a questionnaire for each student from the
original study who was in their class. The instructions for the
completion of the checklist asked the teacher to consider the child
over the last two or three months and for each of the 24 items, write
the number that indicated how true that description was of the child.

The ratings were made on a five-point scale with 1 designated not true,
3 described as somewhat or sometimes true, and 5 as very true.
Teachers were asked to read the items carefully as they were directed
towards several different aspects of the child's behaviour.

The completed questionnaires were returned by post.

RESULTS
The results were analysed initially by exploring principal components
analysis, and subsequently by confirmatory factor analysis and the
Rasch Scaling procedure.

1. Confirmatory Factor Analyses
For the analyses the basic question posed was whether the Student
Behaviour Checklist scale was unidimensional since the
unidimensionality of items is one of the requirements for the use of
the Rasch model (Hambleton and Cook, 1977). If the items were found not
to be unidimensional, it would not be possible to apply the Rasch
procedures to this analysis. Consequently, before doing any kind of
scaling, it was necessary to examine with confirmatory factor analysis
whether the data involved one factor, two factors, a hierarchical or a
nested model. The LISREL8W (Joreskog and Sorbom, 1993) computer program
was used to determine which of the four different models provided the
most adequate explanation of the data collected from the administration
of the Student Behaviour Checklist.

In the graphical representation of the four hypothesised models, the
rectangular boxes indicate the manifest variables, that is, the items
included in the questionnaire, while the ellipse shapes show the latent
variables which are hypothesised to underlie the manifest variables. In
Figure 1 two types of latent variables are shown, that is first order
factors and second order or higher order factor. To differentiate
between the two factor levels, the second or higher factors are
indicated in bold. Arrows in the figure show the direction of influence
from the hypothesised factor to the items to which the teachers
responded. Among the 24 manifest variables, that is the 24 items, items
1, 2, 10, 20, and 24 were selected to illustrate the overall structure
of the confirmatory factor analysis models in which individual items
are assigned to hypothesised factors without showing all the 24
item-factor relationships.

Model 1 is a basic factor model in which manifest variables are
assigned to one single order latent factor, Behaviour. In Model 2 the
24 items in the questionnaire are assigned to either Learned
Helplessness or Mastery in accordance with the specifications of Fincham, Hokoda and Sanders (1989). Items 1, 4, 6, 8, 9, 12, 14, 17, 18, 20, 21, and 23 are learned helplessness items with items 2, 3, 5, 7, 10, 11, 13, 15, 16, 19, 22 and 24 being assigned as mastery items. In this model the errors of measurement are uncorrelated. However, the two factors Learned Helplessness and Mastery are permitted in the analysis to be correlated. Model 3, the hierarchical model is an extension of Model 2 in which it is assumed that the covariance between the first order factors of learned helplessness and mastery is explained by a general higher-order factor of

Figure 1. Hypothetical models of Student Behaviour Checklist Questionnaire Data

Behaviour. When factor models such as Models 2 and 3 are compared with three or less first order factors, the goodness-of-fit of the hierarchical model (Model 3), is identical to that of the two-factor model, that is Model 2 (Marsh and Hocevar, 1985). This implies that it is not possible to produce empirical evidence for the superiority of the hierarchical model, over the two-factor model (Lietz, 1995).

The fourth model is a nested factor model. In this model all the items in the questionnaire are assigned to both one general factor, Behaviour, and to either Learned Helplessness or Mastery. These two factors are correlated but are at the same time orthogonal to the Behaviour factor, which is a general factor. The main difference between the hierarchical and nested factor model is that the nested factor model allows items to be assigned directly to the general factor (Gustafsson and Balke, 1993), while in the hierarchical model the items only contribute to a general factor through the first order factors. All the four models were examined by both a priori and a posteriori analyses in which the errors associated with the manifest variables were allowed to correlated.

(a) a priori analysis
In the a priori analysis the results for the four different models are presented in Table 1. The Chi-square (c^2), degrees of freedom (df), goodness of fit (GFI), adjusted goodness of fit (AGFI), p-value (p), relative noncentrality index (RNI), Tucker Lewis index (TLI) and parsimony noncentrality index (PRNI) indices were taken as a criteria (Swaminathan, 1991; Marsh and Balla, 1994) for comparing the models.

The a priori analysis indicated that the two-factor model (Model 2) and the hierarchical model (Model 3) had the same c^2, df, GFI, AGFI, RNI, TLI and PRNI. It should be noted that in the two-factor model, the latent variables are correlated (r = -0.97) and in the hierarchical
model the factor loadings with Behaviour are 1.00 and -0.73 for Learned Helplessness and Mastery respectively. The p-value in all the four models is zero (0.00) and the GFI is very low. At this stage it was not possible to state that the scale fitted a one factor, two-factor, hierarchical or nested model. Hence, it was necessary to assess the results of the a posteriori analysis of the four models. However, it was evident that the nested model provided the best fit of the four models, even after allowance was made for the reduced degrees of freedom in the adjusted goodness of fit index (AGFI).

(b) a posteriori analysis
In the a posteriori analyses all the possible modifications were carried out to determine the best fitting model among the four alternatives (see Table 1). In these analyses errors associated with the measurement of each item were correlated and in one model, the nested model four items were dropped from the Learned Helplessness and Mastery factors in the model but not the Behaviour factor. The numbers of item correlations and items dropped in each model are also presented in Table 1.

Table 1. The a priori and a posteriori results of the confirmatory factor analyses of the four models

In the a posteriori analysis the one factor and the nested model would appear to be the better models (see Table 1), as their c2, df, p, GFI, AGFI, RNI, TLI and PRNI values approached the levels considered to indicate a satisfactory model. When these two models, were compared the nested model seemed to be the better model, but it should be noted that in the nested model four items were dropped from the lower order structure of the model, but no items were dropped in the one factor model. On this basis, the one factor model would appear to be the best among the four models, indicating that the questionnaire was unidimensional. As the scale clearly met the requirements for unidimensionality proposed by Hambleton and Cook (1977) the Rasch analysis was carried out. Acceptance of the one factor model also indicated that there was no evidence to support the two separate factors Learned Helplessness and Mastery which were hypothesized by Fincham, Hokoda and Sanders (1989) in the development of the instrument. As a consequence of these analyses, it must be argued that the items in the Student Behaviour Checklist measure only one factor Academic Behaviour.

2. Results of Rasch Analyses
After the factor analyses, the second basic question was whether the Rasch model which involves a one factor scale fitted the data. In order to obtain an answer to this question the Rasch analysis was carried out using the QUEST computer program (Adams and Khoo, 1993). Rasch analysis
was chosen as an appropriate procedure for these data, because the Rasch model estimates the person ability independent of the items and the item estimates are independent of the sample. This means that the analysis is a sample and item free procedure. Wright and Stone (1979, 15) explained that
"... the Rasch model allows us to estimate person ability and item difficulty independently of one another in such a way that the estimates of person ability are freed from the effects of the item difficulty and the estimates of item difficulty are freed from the effects of person ability".

Moreover, the items and the persons are brought to a common scale.

A choice was made among the different Rasch procedures to employ rating scale analysis in this study. The Rating scale procedure was selected, because it "assumes a single underlying dimension for the variable and seeks to scale the data in such a way that interval scale data are obtained for the variable formed" (Wolf, 1994, 4926). The responses, however, involved unipolar scales with the same response categories across all items. Rating scale analysis was also the preferred technique for the analysis of such response categories since the questionnaire in this study required the judgement of teachers about the behaviour of their students. In situations, when human judgement is used, a rating scale is the best means of securing these judgments (Wolf, 1994).

The results of the analysis are shown in Table 2.

At the beginning of this study an exploratory principal components factor analysis using the SPSS computer program was carried out to examine the factor loadings on Learned Helplessness and Mastery. The results indicated that all the Mastery items had negative factor loadings while all the Learned Helplessness items were positively loaded. That is the Mastery and Learned Helplessness items were loading in opposite directions. Consequently the principal components and confirmatory factor analyses both indicated that it was necessary to reverse the Learned Helplessness items responses from (01234) to (43210) during the rating scale analysis.

Initially the whole scale of 24 items was analysed using the QUEST computer program (Adams and Khoo, 1993). In the first analysis it was found there were 14 misfitting items (see Table 2). The infit mean squares of these misfitting items were outside the acceptable range of 0.83 and 1.20. In Rasch analysis, items that do not fit the Rasch model must be deleted from the scale (Rentz and Bashaw, 1975; Wright and Stone, 1979; Kolen and Whitney, 1981; Smith and Kramer, 1992). Hence the misfitting items were deleted one at a time. If there are many misfitting items in any one analysis it is important to choose which one item to delete in the next analysis. Item 11 was the first item to
be deleted as its infit mean square was 2.37. After 14 items, which were considered to be misfitting by this criterion, had been deleted on this basis only ten items which fitted the Rasch scale remained (see Table 2).

Weiss and Yoes (1991) have suggested that there must be a truce between the discrimination or the total information accommodated (fidelity) by an item and the range (bandwidth) over which that information is available. Table 2 indicates that even if 14 of the items did not fit the Rasch scale, the overall discrimination power of the items was very high. Items which have higher discrimination power such as items 3, 7, 15, 17. have a high fidelity but a narrow bandwidth. Such items with high discrimination only provide information over a narrow ability range and little or no information outside that range (Weiss and Yoes, 1991).

Table 2. Results of Rasch analyses

<table>
<thead>
<tr>
<th>ITEMS</th>
<th>Before Deletion</th>
<th>After Deletion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 item 1</td>
<td>1.02</td>
<td>0.70</td>
</tr>
<tr>
<td>2 item 2</td>
<td>0.67</td>
<td>Deleted</td>
</tr>
<tr>
<td>3 item 3</td>
<td>0.53</td>
<td>Deleted</td>
</tr>
<tr>
<td>4 item 4</td>
<td>1.19</td>
<td>0.68</td>
</tr>
<tr>
<td>5 item 5</td>
<td>0.75</td>
<td>Deleted</td>
</tr>
<tr>
<td>6 item 6</td>
<td>0.94</td>
<td>0.70</td>
</tr>
<tr>
<td>7 item 7</td>
<td>0.90</td>
<td>0.77</td>
</tr>
<tr>
<td>8 item 8</td>
<td>1.31</td>
<td>Deleted</td>
</tr>
<tr>
<td>9 item 9</td>
<td>0.96</td>
<td>0.69</td>
</tr>
<tr>
<td>10 item 10</td>
<td>0.89</td>
<td>Deleted</td>
</tr>
<tr>
<td>11 item 11</td>
<td>2.37</td>
<td>Deleted</td>
</tr>
<tr>
<td>12 item 12</td>
<td>0.66</td>
<td>Deleted</td>
</tr>
<tr>
<td>13 item 13</td>
<td>0.93</td>
<td>0.74</td>
</tr>
<tr>
<td>14 item 14</td>
<td>1.47</td>
<td>Deleted</td>
</tr>
</tbody>
</table>
15 item 15 0.59 m 0.79 Deleted
16 item 16 1.29 m 0.47 Deleted
17 item 17 0.64 m 0.79 Deleted
18 item 18 0.82 m 0.73 0.870.72 -0.05
19 item 19 0.69 m 0.73 Deleted
20 item 20 1.08 0.671.060.690.10

21 item 21 1.56 m 0.50 Deleted
22 item 22 0.96 0.660.990.66 -0.02
23 item 23 0.77 m 0.77 Deleted
24 item 24 0.91 0.660.950.650.03

================================================================

m Misfitting items outside the accepted range of 0.83 to 1.20

Items with low discrimination power such as items 8, 11, 14, 16 and 21 provide information over a wide ability range and have a wide bandwidth but a low fidelity. Because of the lack of balance between bandwidth and fidelity (see Table 2) many of the highly discriminating items did not fit the Rasch scale. In total 14 items had to be deleted.

After the deletion of the 14 items the distribution of the cases and the threshold levels of the ten items are shown in Figure 2. The mean of the item thresholds was set at 0.00 and the standard deviation was 0.28, while the mean of the cases was 0.77 and the standard deviation was 1.10. The standard deviation is almost one logit of this scale, with the logit being the natural unit of the Rasch scale (Bear and Pettie, 1979). In all, 198 students were at or above the item mean. It should be noted that a peak occurs on the frequency distribution of cases at about -0.80 after a hollow at -0.70, and of the 60 students below the item mean 17 students who were between -0.7 and -1.00 logits could be identified as demonstrating marginal academic behaviour while there are other six students who score below -1.00 and are clearly demonstrating marked behavioural problems.

With respect to the items, only ten out of 24 items fitted the Rasch scale. Of the ten items that fitted the Rasch scale, six were learned helplessness items and four were mastery items (see Table 3). Although the scale was constructed to measure both Learned Helplessness and Mastery and more importantly to allow teachers to identify students exhibiting these characteristics, both the confirmatory factor analysis
and the Rasch analysis would suggest that the checklist operates as a single scale and measures a characteristic which may be referred to as academic behaviour.

The items in the final scale relate to effort (items 1 (LH) and 13 (MO)), motivation (items 4 (LH) and 7 (MO)), reaction to failure (items 6 (LH), 9 (LH) and 24 (MO)), persistence (items 20 (LH) and 22(MO)), and response to teacher inquiry (item 18 (LH)). These items relate clearly to indices of academic behaviour. This modified scale allows teachers to discriminate between students, as well as allowing for the identification of students with respect to academic behaviour, with behavioural difficulties on a scale of measurement that is independent of the items employed and the students in the sample, once the zero of the scale has been set.

Figure 2.
Rating Scale STUDENT BEHAVIOUR CHECKLIST(24 ITEMS) Item Estimates (Thresholds)
------------------------------------------------------------------------
-----------------
all on behaviour (N = 258 L = 10 Probability Level=0.50)
------------------------------------------------------------------------

| 4.0 |
|-----------------
|   XXXXXXXXXXXX |
|          X |
| 3.0 |
|   XXXXXXX |
|          X |
|   XXXXXXXXXXXXXXXXXXX |
| 2.0 |
|   XXXXX |
|           1.4 |
|   XXXXXXXXXXXX |
|              13.4 |
|   XXXXXXXXXXXX |
|             20.4 24.4 |
|   XXXXXXX |
|            6.4 7.4 18.4 22.4 |
|   XXXXXXXXXXXX |
|             4.4 |
|   XXXXXXX |
| 1.0  XXXXXXXXXXXX |

...
Table 3 The Academic Behaviour Scale: A modification of the Student Behaviour Checklist

<table>
<thead>
<tr>
<th>CharacteristicsLearned Helplessness items (LH)</th>
<th>Mastery Oriented items (MO)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effort</td>
<td></td>
</tr>
<tr>
<td>1 Prefers to do easy problems rather than hard ones.</td>
<td></td>
</tr>
</tbody>
</table>
13. Prefers new and challenging problems over easy problems.

Motivation
4. Takes little independent initiative; you must help him/her to get started and keep going on an assignment.

7. Tries to finish assignments, even when they are difficult.

Failure
6. When s/he fails one part of a task, s/he looks discouraged—says s/he is certain to fail at the entire task.

9. Gives up when you correct him/her or find a mistake in his/her work.

24. When s/he receives a poor grade, says s/he will try harder in that subject the next time.

Persistence
20. Says things like "I can't do it" when s/he has trouble with his/her work.

22. When experiencing difficulty, s/he persists for a while before asking for help.

Response to teacher inquiry
18. Does not respond with enthusiasm and pride when asked how s/he is doing on an academic task.

Discussion

This study set out to examine teachers' perceptions of student behaviour and the psychometric properties of the Student Behavior Checklist. The central method employed in the analysis of the data was the use of Rasch scaling procedures, but this required the use of confirmatory factor analysis.

The findings with respect to the psychometric properties of the Student
Behavior Checklist were initially surprising, as rather than finding evidence for separate scales of learned helplessness and mastery orientation, both the confirmatory factor analysis and the Rasch analysis clearly identified the scale as being unidimensional. Further the Rasch rating scale analysis indicated that of the 24 items only 10 fitted the measurement scale. However because Rasch analysis enables the scale to be established independently of the sample, and the characteristics of the sample to be examined independently of the scale, the ten items forming the measurement scale constitute a sound psychometric instrument that has the properties of an interval scale. When the item estimates for the ten items were examined and plotted on a map (see Figure 2) it was clear that teachers had made distinctions between students and in particular that while most students were above the scale zero or mean of the item threshold, some students at the lower end of the scale could be identified as having behavioural difficulties.

When the 10 acceptable items were examined it was found that 6 of these had been designated by Fincham et al (1989) as measuring learned helplessness and four as measuring mastery orientation. Yet it was apparent that rather than the teachers seeing these as separate entities, they indicated that these items measured only single type of behaviour. An alternative explanation may be that there is in fact no discernible difference between learned helplessness and mastery orientation which could be detected by teachers within classrooms. Interestingly when the 10 items (see Table 3) are examined with respect to the criteria for learned helplessness suggested by Peterson et al (1992), Item 1 clearly relates to a reduction in behavioural agency, with Item 13 as its antithesis, Item 4 relates to motivation with Item 7 as its antithesis, and Items 6 and 9 relate to changes in cognition and emotion. This reaction to failure aspect measured in Items 6 and 9 is countered by Item 24 which measures an increase or renewal of effort in the face of failure. In addition, Item 10 relates to lack of enthusiasm and pride in response to teacher enquiry. Interestingly, this trait has also been reported by Yates et al (1995) as being a significant difference between pessimistic and optimistic children in relation to their reported attitudes towards mathematics.

The variability of teacher judgements noted in the reviews of the literature by Hoge and Coladarci (1989) and Brookhart (1994) is not apparent in many of the items that were deleted, because these items had high discrimination indices and their band widths were very narrow indicating that the teacher ratings on these items provided information over a very limited range. However the extent to which the teacher ratings were valid is as yet unanswered by this study. Future work should involve an examination of the relationship between teacher ratings on the modified 10-item Student Behaviour Checklist, measures of student achievement and student self reports. Such analyses will be
possible when the data currently being collected becomes available.

The question as to whether the scale measures learned helplessness and mastery is clearly answered. Within teachers judgements these traits do not constitute separate characteristics. Fincham et al (1989) suggested that the scale may be measuring academic competence - certainly the ten items can be conceptualised as constituting a scale of academic behaviour, with six designated learned helplessness items clearly relating to a lack of academic behaviour and the designated mastery orientation items relating to the presence of academic behaviour. It should be noted that teachers are providing ratings of overt behaviour, rather than inferring students' internal states. Spivak and Swift (1973) noted that when asked to rate overt behaviours teachers do discriminate between groups, with their ratings being stable over time. Certainly the stability of teacher ratings could be further investigated, with the modified scale. So too could the variability in rating between teachers. In addition, the cluster analysis of grading practices of teachers from which Frary et al (1993) identified teachers as using either a norm referenced approach or as being soft hearted, strict, arbitrary, uncertain or inconsistent, could be investigated with respect to teacher ratings of student behaviour.

In developing the Student Behavior Checklist, Fincham et al (1989) highlighted the need both for a shorter version of the scale and to tap teacher perceptions as a means of either supplementing or replacing student self report measures. This modified scale of ten items certainly meets the first need. Interestingly, Fincham et al (1989) note that most studies have focussed on students in Grade 5, since this is the group most amenable to pencil and paper self reporting. Younger children are more difficult to assess with such measures. However this modified scale of academic orientation should be cost effective in use, is composed of only 10 items, and is completed easily by the teacher. Moreover, since it is robust psychometrically it could perhaps be used to identify academic disinterest in younger children. Certainly, Fincham et al (1989) express the expectation that since such behaviour would seem to relate to a pattern that retards learning, particularly when the learning involves material that is difficult for the child, its early detection would be advantageous.

Conclusion

1 The Student Behavior Checklist is a unidimensional scale of academic behaviour.
2 Support was not found for separate scales of learned helplessness and mastery orientation.
3 Of the 24 items, only 10 fitted the Rasch model.
4 These ten items constituted a psychometrically robust instrument that has the properties of an interval scale.
With these 10 items forming a scale that measures academic orientation, teachers could detect behavioural differences between students.

The use of Rasch analysis allowed for the detailed examination of the Student Behavior Checklist, thus significantly advancing knowledge of teacher perceptions of student behaviour.

REFERENCES


factor analysis to the study of self-concept: First- and higher-order factor models and their invariance across groups. Psychological Bulletin, 97, 562-582.

Educational Psychology, 15, 23-35