Edith Cowan University
Research Online

ECU Publications Pre. 2011

2009

Visual Discrimination Of Letters In The Alphabet By Young Children: A Rasch Measurement Analysis

Janet Richmond
Edith Cowan University

Russell Waugh
Edith Cowan University

This article was originally published as:
Richmond, J. E., & Waugh, R. F. (2009). Visual discrimination of letters in the alphabet by young children: A Rasch measurement analysis. *Proceedings of AARE Canberra Conference 2009*. (pp. 1-28). Canberra. AARE.
This Conference Proceeding is posted at Research Online.
http://ro.ecu.edu.au/ecuworks/6546

RIC09924

Visual Discrimination of Letters in the Alphabet by Young Children:

A Rasch Measurement Analysis

Janet Richmond and Russell Waugh

Faculty of Education and Arts, Edith Cowan University

Perth, Western Australia

Key words: visual discrimination, reading, alphabet letters, Rasch measurement, primary, students

A paper to be presented at AARE 30 November to 3 December 2009 in Canberra

Address correspondence to: Professor Russell F. Waugh, School of Education, Edith Cowan University, Bradford Street, Mount Lawley, 6050, Western Australia.

Telephone (08) 9370 6941
r.waugh@ecu.edu.au

ABSTRACT

Data on 30 items for upper case letters and 36 items for lower case letters, where each item was scored in one of two categories (wrong scored zero and correct scored one), were Rasch analysed to create two linear scales . The student sample was N=324 pre-primary and primary students (aged 5 to 9 years) in Perth, Western Australia. Twelve of the initial 30 items of Visual Discrimination of Upper Case Letters were deleted due to item misfit statistics leaving 18 items and five of the original 36 items for Visual Discrimination of Lower Case Letters were deleted leaving 31 items. The 18 item-scale and the 31 item-scale each had a good fit to the measurement model, were reliable (Person Separation Indices of 0.55 and 0.82 and Cronbach Alphas of 0.70 and 0.82), and were unidimensional, showing no statistically significant interaction on item difficulties along the scale. Items were ordered from easy to hard and student measures from low to high on the same scale, allowing the objective identification of alphabet letters that students found difficult to discriminate. Students who had poor visual discrimination skills of alphabet letters could also be identified objectively.

Visual Discrimination of Letters in the Alphabet by Young Children: A Rasch Measurement Analysis

Introduction

For satisfactory academic development it is expected that children perform adequately for the age or grade level of a child in the areas of reading, spelling, writing, mathematical computations, communicating, science, computers, sports, among other areas of academic performance (Erhardt & Duckman, 2005; Kirk, Gallagher, & Anastasiow, 2000; Loikith, 1997). Concepts required for these academic skills such as the concept of space, are dependant to a great extent on visual perception and are reflected in language in words that explain size, shape, colour, number, position, direction and distance. Comprehension of these words and concepts in, for example, listening, reading, mathematics and geography reflect the adequacy of the visual spatial functions of the individual (Carrow-Woolfolk, 1981). Some authors state that letter discrimination (the ability to see the visual differences between letters) and letter identification processes (knowledge of the correspondence between letters and phonemes), as well as visual attention and memory, are involved in reading (Catts & Kamhi, 1999; Schneck, 1996). These theories determine that accurate, effortless word recognition requires the use of visual decoding based on familiar letter sequences or graphic configuration and orthographic patterns (order of letters), while phonological skills (sounds represented) are necessary to develop proficient word recognition and semantics (meaning) (Catts & Kamhi, 1999; Goldstand, Koslowe, & Parush, 2005; Kulp, 1999; Lachmann & Geyer, 2003). Thus, weak readers demonstrate their visual perceptual difficulties related to spatial concepts and shape recognition by their slow reading speed, since they have to sound out each word and confuse letters such as “b” and “d” (Green & Chee, 1997). The underlying causes of reading problems may differ between beginning readers and poor readers, as well as between poor readers depending on the pattern of reading performance (Lachmann & Geyer, 2003; Oliver, Dale, & Plomin, 2007; Schneck, 2005). These authors found that reversals did not predict the performance on reading tests in young children, but were a good predictor of performance on reading tests for grade 3 children, possibly due to their age and developmental level. Children who display more difficulties discriminating orientationally-related letters or patterns show more reversals in reading text (Lachmann & Geyer, 2003). It is thus important that children with potential difficulties in reading are identified early by using linear scales to assess their letter discrimination ability.

The current study was aimed at producing a linear scale which would measure the ability of primary school children to visually discriminate upper and lower case letters in readiness for learning to read.

Current Measures of Visual Discrimination

Identifying students with visual discrimination problems and identifying objectively those letters that students have difficulty discriminating is an important issue in the teaching of reading to young children and in finding out how to help them read better (see for example Wolf, 2008). Current instruments used to assess visual perceptual aspects of letters and letter reversal recognition skills include the *Jordan Left-Right Reversals Test* (JLRRT) (Jordan, 1990), the *Reversal Frequency Test* (RFT) (R. A. Gardner, 1978) and the *Test of Pictures, Forms, Letters, Numbers, Spatial Orientation and Sequencing Skills* (TPFLNSOSS) (M. F. Gardner, 1991). The goal of these assessments is to discover what knowledge children bring to the visual task, in other words, the children's abilities, strengths and weaknesses (Gregg & Scott, 2000; Loikith, 1997). It is important that the instruments used to measure these letter discriminations and letter reversals actually measure what they purport to measure objectively and accurately (Bailey, 1991; Clegg, 1982; Cooke, McKenna, Fleming, & Darnel, 2006; Downing, 2003; McDaniel, 1994; Messick, 1995a, 1995b).

According to Richardson (1996), paediatric occupational and speech therapists are increasingly making use of standardised tests to determine eligibility for therapy reading services, monitor progress and decide about the type of treatment required. Standardised tests allow for the measurement of the child's performance in a specific area according to the 'norm' or average for a particular age level. However, existing standardised tests were developed using True Score Theory Measurement and the measures are not linear. True Score Theory says that a total raw score on a test equals the 'true score' (which cannot be observed) plus a random error score and almost any test data will fit this measurement model. In addition, the reliability and validity of the current tests are called into question.

The TPFLNSOSS and the RFT do not require a verbal response or reading/language comprehension and, in Level Two of the JLRRT, a degree of reading comprehension is required for successful completion of the test. The TPFLNSOSS combines visual perception with classroom related tasks, while the JLRRT considers reversals of letters, numbers, letters in words, as well as whole word reversal. All the tests have gaps in their psychometric evidence (either not presented or low reliability coefficients, for example). The RFT does not report any psychometric data and fails to adequately explain the rationale, has a poorly written manual with little detail and some ambiguity and vagueness. The TPFLNSOSS reaches a ceiling where there are not enough difficult items for seven, eight and nine year olds and the paper is of a poor quality allowing the print to show through the page resulting in possible confusion to the child. The JLRRT reports an inflated reliability for the older child due to the low development of skills. Performance on Level Two is strongly related to reading and comprehension ability. The assessments available to discriminate letters and letter reversals visually display flaws in their development. These flaws and inconsistencies led to the conclusion that a new assessment of visual letter and letter reversals had to be developed which had linear measures and was psychometrically sound to enable valid inferences to be made. This would ensure that children with letter and number reversal problems could be accurately identified and remedial strategies could be started at an early stage.

Methodology

The two tests reported here were part of a larger study developed for the measure of letter and number, and letter and number reversal recognition (discrimination) but only the Rasch analysis for upper case and lower case letters are reported here. Ethical and administrative approvals were obtained from the university and from the schools, parents and students, with signed informed consent forms. In the Visual Discrimination of Upper Case Letters Test, the students were presented with upper case letters in a random order where some letters are reversed and some letters are facing the right way. In all cases where a letter is not symmetrical around the vertical axis, both a reversed and correctly oriented letter was presented in random order and not in close proximity. Each letter was spaced apart from the next, so that it is easy to isolate each letter. The students were required to indicate which of the upper case letters are reversed on the page. Visual Discrimination of Lower Case Letters consists of a similar random presentation of lower case letters where all the letters are presented in the correct direction and those that are not symmetrical around the vertical axis are also presented in the reversed orientation. The students were requested to indicate the reversed lower case letters. The items were discussed with several occupational therapists, speech pathologists and reading experts who helped in deciding which items to use. The final sets of items determined after the Rasch analysis are given in Tables 5 and 6.

A pilot study was conducted with 20 students aged 5 to 10 years and was followed by an interview with each student. The tests were all completed during one single session lasting between 30 and 50 minutes, depending on the students work pace and the age of the students. After the students had completed the tests, they were asked to verbally feedback on their subjective experience and opinion of the test by answering four questions: Would you mind telling me if you found the test interesting? What made it interesting/boring? What do you think should be done differently? How did doing the test make you feel? The child's responses were recorded by the examiner on an interview record sheet.

All the Pre-primary and Year One students reported that the original battery of tests was too long and an attempt was made to shorten the tests, after further discussion with the reading experts. The year Two and year three students had varying opinions about the length of the tests with some saying it was the right length and one saying it was too long. Comments received about what made the eight tests interesting were the novelty of the tests, the challenge to complete the tests without making any mistakes or not getting "caught by the tricks". The only comment about what could be done differently was that the tests should be shorter. Most students said they felt "OK" about doing the test. Some students said they did not like the test because it was difficult or they found the tests too lengthy, three of the students enjoyed completing the tests and requested to do it again.

Seven primary Schools in and around the Perth metropolitan area in Western Australia were used for the data collection. The data collection occurred over a three month period from October to December 2008. Students were included in the study if they were between the ages of five to ten years old, volunteered and signed consent forms were obtained. An information and consent form was sent home with every Pre-Primary to Year Three student in the seven schools with a request that the form be returned to the class teacher by a set date. Students were required to have a working knowledge of the English language to complete the assessment. Students with known developmental disorders, intellectual limitations, neurological impairments, learning difficulties, psychiatric disorder and/or visual difficulties, as identified on the parent report form, were not excluded from the participant group.

A convenience sample of 324 students was acquired, ensuring the inclusion of five public primary schools, and two independent schools by subdividing the schools into categories prior to the selection process. These participants formed the sample for the main data collection. Every child from Pre-primary to Year Three in the participating schools was given an opportunity to participate in the study. The return rate of the parent consent forms was between 10% and 30% from the various schools. School Principals reported that this was a generally accepted return rate for any forms sent out by the schools. The sample included 177 girls and 146 boys. There were 45 Pre-primary students, 118 Year One students, 77 Year Two students and 83 Year Three students. Twenty-nine of the students were four or five years old, 71 were six years old, 92 were seven years old, 87 were eight years old, 39 were nine years old and six were ten years of age. Seventy-two students were reported by the parents as having had some form of intervention or diagnosis relating to learning difficulties, while 252 students had no record of previous or current interventions or learning difficulty. There were 68 students who attended private schools, while 256 students attended public schools.

Initial Rasch Analysis

An initial Rasch analysis was performed on the original items for Visual Discrimination of Upper Case Letters (30 items) and Visual Discrimination of Lower Case Letters (36 items) where each item was scored in one of two categories (wrong scored zero and correct scored one). The computer program used was Rasch Unidimensional Measurement Models (RUMM2020) (Andrich, Sheridan & Luo, 2005). Twelve of the initial 30 items of Visual Discrimination of Upper Case Letters were deleted due to item misfit statistics, meaning that half of students scored the item correct and half of the students scored the item as incorrect and the item was therefore considered to be non-discriminatory. The disagreement between the students may be related to the chosen Modern Victorian Font which was used in the scale as this is the script that is taught to students in some states in Australia. The remaining 18 items were found to have a reasonable fit to the measurement model for the 324 students included in this study. For Visual Discrimination of Lower Case Letters, five of the initial 36 items were deleted due to item misfit statistics, where students were divided about the difficulty of the items and they were therefore non-discriminatory, possibly also due to the Modern Victorian Font style used in the scale. The remaining 31 items displayed a good fit to the measurement model. The Rasch analysis with the RUMM program does not indicate how to alter an item in order to make it fit the measurement model. In order to include, in a future measure, the deleted items which were initially considered conceptually valid, these would need to be changed and re-tested. One suggestion, from anecdotal evidence, is to change the font used in the scale to something with which the students might be more familiar in printed context.

Final Rasch Analysis Results

The following output shows the results for the final Rasch analysis for Visual Discrimination of Upper Case Letters (18 items) and Visual Discrimination of Lower Case Letters (31 items) with the RUMM2020 computer program.

Summary of Fit Statistics

The RUMM2020 program estimates an item-person interaction which establishes the overall fit statistics that determine whether the item estimations contribute meaningfully to the measurement of one construct. This calculation thus examines the consistency with which students' responses agree with the calculated difficulty of each item on the scale. The standardised fit residual statistics (see Table 1) have a distribution with a mean near zero and a standard deviation near one when the data fit the measurement model (Andrich, 1985), as is the case with these three measures. This means too that there is a good pattern of person and item responses consistent with a Rasch measurement model.

Dimensionality

For Visual Discrimination of Upper Case Letters, there was an item-trait interaction chi-square of 42.07 with df = 36 and a probability of 0.23. This means that the scale is constructed with reasonable agreement amongst the students about the linear progressive difficulty of the items. The item-trait interaction chi-square for Visual Discrimination of Lower Case Letters was 136.85 with df = 124 and a probability of 0.20, showing a similar reasonable agreement amongst the students about the linear progressive difficulty of the items along the scale.

Table 1: Global Item and Student Fit Residual Statistics (N=324)

	ITEMS		PERSONS	
	Location	Fit Residual	Location	Fit Residual
Visual Discrimination of Upper Case Letters (I=18)				
Mean	0.00	-0.70	2.99	-0.44
Standard Dev.	0.77	1.36	0.81	0.79
Visual Discrimination of Lower Case Letters (I=31)				
Mean	0.00	-0.50	2.68	-0.56
Standard Dev.	1.21	1.55	1.31	1.02

Comment on Table 1
Fit residuals have a mean near zero and a standard deviation near one when the data fit the measurement model (as is the case here). This reflects good consistency of item and student scoring patterns.

Person Separation Index

The Person Separation Index is an estimate of the true score variance among the students and the estimated observed score variance using the estimates of their ability measures and the standard error of these measures (Andrich & van Schoubroeck, 1989). For Visual Discrimination of Upper Case Letters and Lower Case Letters, the Person Separation Indices are 0.55 and 0.82. Cronbach Alphas are 0.70 and 0.82 and so internal reliability is also satisfactory. For a good measure, it is desirable that these indices should be 0.9 or greater, as it is an indicator that the student measures are separated by more than their standard errors. Based on this index, the Visual Discrimination Lower Case Letters scale demonstrates acceptable separation, but Visual Discrimination Upper Case Letters requires improvements to the measure in any future use.

Individual Item Fit

Items are ordered by calibrated values to evaluate their fit to the measurement model. The location of each item on the scale is the item difficulty in standard units, called logits (log odds of answering successfully). All the items fit the measurement model with probabilities greater than p=0.10 (see Table 2). The residuals shown in Table 2 represent the difference between the observed responses and the expected responses calculated from the Rasch measurement parameters. Standardised residuals should fall within the range of -2 and +2. Table 2 shows that all items for Visual Discrimination Upper Case Letters have acceptable residuals except for item 30.

Table 2: Individual Item Fit Statistics for Visual Discrimination Upper Case Letters

Item	Location	SE	Residual	DegFree	ChiSq	DegFree	Prob
2	-1.58	0.51	-1.99	136.94	3.35	2	0.19
24	-0.92	0.39	-1.40	136.94	2.12	2	0.35
27	-0.57	0.34	-0.86	136.94	1.81	2	0.40
18	-0.56	0.34	-1.04	136.94	1.26	2	0.53
25	-0.43	0.32	-1.48	136.94	1.69	2	0.43
5	-0.35	0.31	-1.17	136.94	2.21	2	0.33
1	-0.32	0.31	-1.42	136.94	4.45	2	0.11
22	-0.18	0.29	-0.31	136.94	2.69	2	0.26
28	-0.18	0.29	-0.99	136.94	0.97	2	0.62
17	0.02	0.27	-1.14	136.94	2.89	2	0.24
11	0.11	0.27	-0.52	136.94	1.36	2	0.51
15	0.17	0.26	-1.71	136.94	3.27	2	0.19
23	0.31	0.25	-0.88	136.94	1.62	2	0.44
4	0.32	0.25	-1.74	136.94	3.33	2	0.19
3	0.39	0.24	-0.50	136.94	4.53	2	0.10
9	0.66	0.22	-0.67	136.94	1.37	2	0.50
8	1.29	0.19	1.44	136.94	1.15	2	0.56
30	1.79	0.18	3.84	136.94	1.99	2	0.37

For Visual Discrimination of Lower Case Letters, all the items fit the measurement model with probabilities greater than p=0.08 (see Table 3), but a few of the residuals are a little outside what might be considered ideal limits.

Table 3: Individual Item Fit Statistics for Visual Discrimination Lower Case Letters

Item	Location	SE	Residual	DegFree	ChiSq	DegFree	Prob
9	-2.06	0.42	-2.47	248.71	4.81	4	0.31
8	-1.83	0.38	-0.61	248.71	4.05	4	0.40
6	-1.51	0.33	-1.97	248.71	4.69	4	0.32
10	-1.30	0.31	-0.22	248.71	4.16	4	0.39
29	-1.19	0.29	-2.46	248.71	3.49	4	0.48
1	-1.15	0.29	-1.72	248.71	2.39	4	0.66
13	-1.10	0.28	-2.05	248.71	2.14	4	0.71
16	-1.08	0.28	-2.74	248.71	5.91	4	0.21
19	-1.06	0.28	-0.98	248.71	2.17	4	0.71
34	-0.90	0.26	-1.09	248.71	1.50	4	0.83
31	-0.70	0.25	-0.21	248.71	3.04	4	0.55
28	-0.54	0.23	-0.90	248.71	6.39	4	0.17
17	-0.49	0.23	-0.85	248.71	3.32	4	0.51
21	-0.35	0.22	-0.82	248.71	4.98	4	0.29
35	-0.21	0.21	0.16	248.71	6.37	4	0.17
25	-0.11	0.20	-2.23	248.71	5.46	4	0.24
23	-0.11	0.20	-2.82	248.71	7.06	4	0.13
30	0.11	0.19	-0.52	248.71	3.41	4	0.49
3	0.12	0.19	-0.76	248.71	2.34	4	0.67
2	0.26	0.18	-0.91	248.71	6.30	4	0.18
26	0.34	0.18	-1.71	248.71	2.14	4	0.58
7	0.81	0.16	-0.81	248.71	7.80	4	0.10
14	0.99	0.16	0.21	248.71	6.76	4	0.15
15	1.12	0.15	0.79	248.71	5.81	4	0.21
33	1.28	0.15	2.03	248.71	4.18	4	0.38
12	1.46	0.15	1.27	248.71	2.48	4	0.65
24	1.59	0.14	1.08	248.71	3.94	4	0.41
31	1.62	0.14	1.27	248.71	3.56	4	0.47
36	1.78	0.14	1.61	248.71	2.24	4	0.69
11	1.93	0.14	1.41	248.71	5.02	4	0.29
4	2.26	0.14	3.47	248.71	8.25	4	0.08

Notes on Table 2 and 3

1. Location refers to the difficulty of the item on the linear scale.
2. SE means Standard Error, and refers to the degree of uncertainty in a value.
3. Residual represents the difference between the expected value of an item, calculated according to the Rash measurement model and the actual value.
4. DegFree stands for degrees of freedom, and refers to the number of scores in a distribution that are free to change without changing the mean distribution.
5. ChSq stands for Chi-square
6. Prob relates to the probability based on the Chi-square and refers to the levels of certainty to which an item fits the measurement model.

Targeting

The RUMM2020 program produces a student-measure item-difficulty or targeting graph on which the student measures are placed on the same scale as the item difficulties in standard units called logits. For Visual Discrimination of Upper Case Letters (see Figure 1), this targeting graph shows that the student measures cover a range of about -0.8 to +3.5 logits and the item difficulties cover a range of about -1.5 to +1.8 logits. From the graph it can be seen that many students (about 290) were able to answer the items correctly and the targeting of the items needs to be improved in any future use of the scale by adding in some harder items to 'cover' the students with the higher measures.

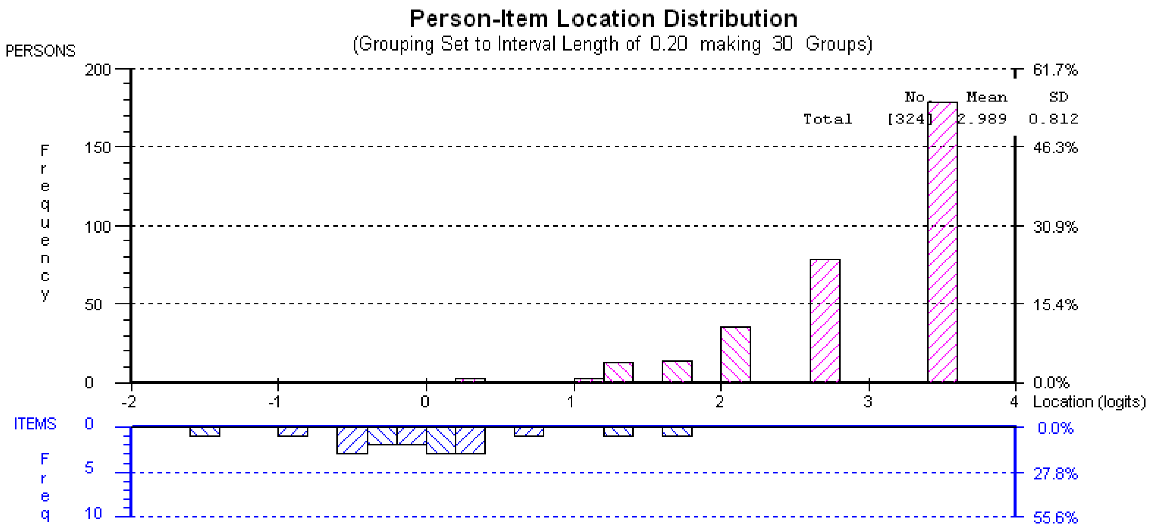


Figure 1 Targeting Graph for Visual Discrimination Upper Case Letters
Note: Student measures are on the upper side in logits. Item difficulties are on the lower side of the same scale in logits. Many students (about 290) answered the items correctly.

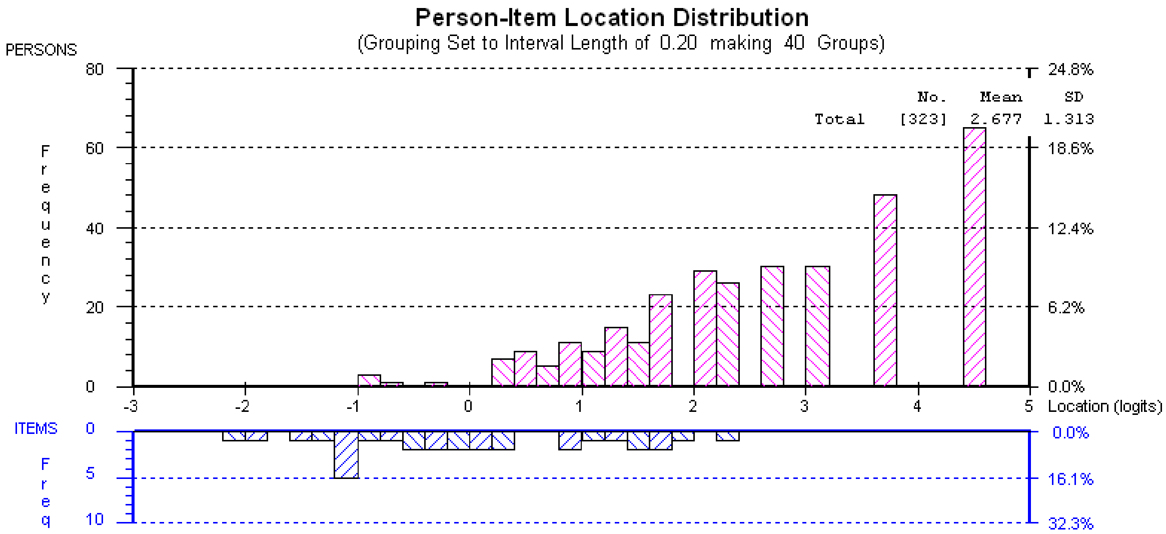


Figure 2: Targeting Graph for Visual Discrimination Lower Case Letters
Note: Student measures are on the upper side in logits. Item difficulties are on the lower side of the same scale in logits. Many students (about 175) answered the items correctly.

For Visual Discrimination of Lower Case Letters (see Figure 2), the targeting graph shows that the student measures cover a range of about -1.0 to +4.5 logits and the item difficulties cover a range of about -2.2 to +2.3 logits. From the graph it can be seen that many students (about 175) were able to answer the items correctly and the targeting of the items needs to be improved in any future use of the scale by adding in some harder items to 'cover' the students with the higher measures.

Discrimination

Item Characteristic Curves examine the relationship between the expected response and the mean group student measures. These curves display how well the item discriminates between groups of persons. An example of one item characteristic curve for each construct will be presented. Figure 3 shows the Item Characteristic Curve for Item 26 Visual Discrimination of Lower Case Letters. This curve shows that the item discriminates well for students with different measures. The Item Characteristic Curves for all the other items were checked and found to be satisfactory (but are not reported here to avoid unnecessary repetition).

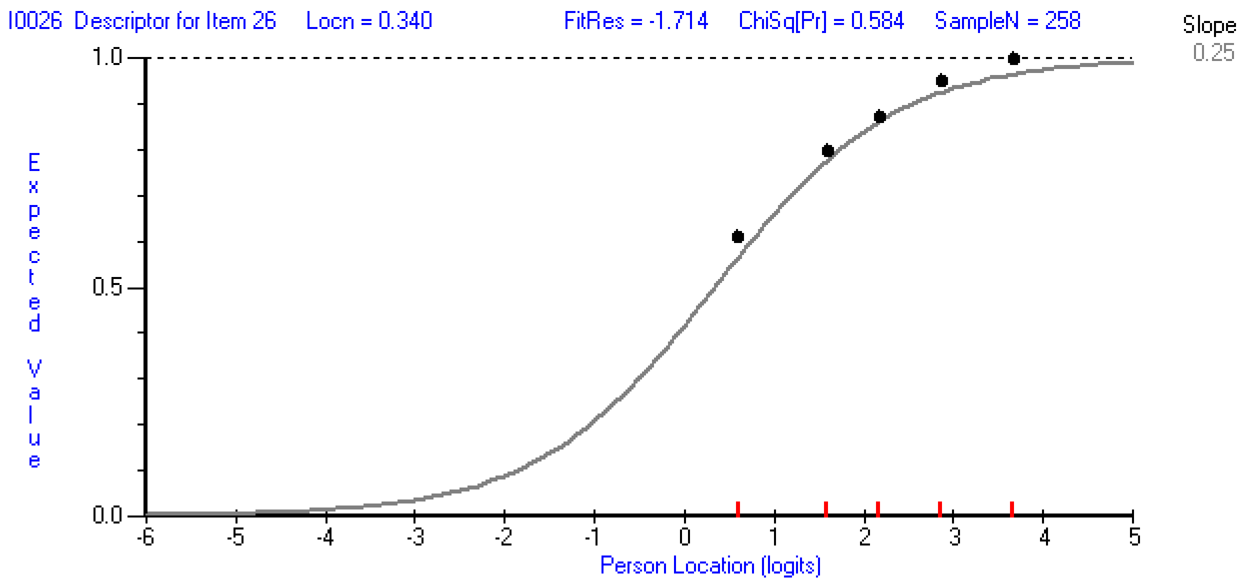


Figure 3: Item Characteristic Curve: Item 28-Visual Discrimination Lower Case Letters

Consistency of Use of Scoring Categories

The RUMM2020 program produces graphs of the scoring categories for each item. The Scoring Category Curves show the relationship between the probability of scoring in each category (zero for wrong and one for right) on each item. Figure 4 is the Scoring Category Curve for item 1 of Visual Discrimination Upper Case Letters. This figure shows that the scoring was done logically and consistently. When students have low measures on item 1, then they have a high probability of obtaining a zero score (answer wrong) and, when they have a high measure, they have a high probability of scoring 1 (answer correct). The Scoring Category Curves for all the other items were checked and they were satisfactory too. The Scoring Category Curves for all the items of the other two variables, Visual Discrimination Lower Case Letters and Visual Discrimination Numbers, were checked and they were also found to be satisfactory, but they are not presented here to avoid too much repetition.

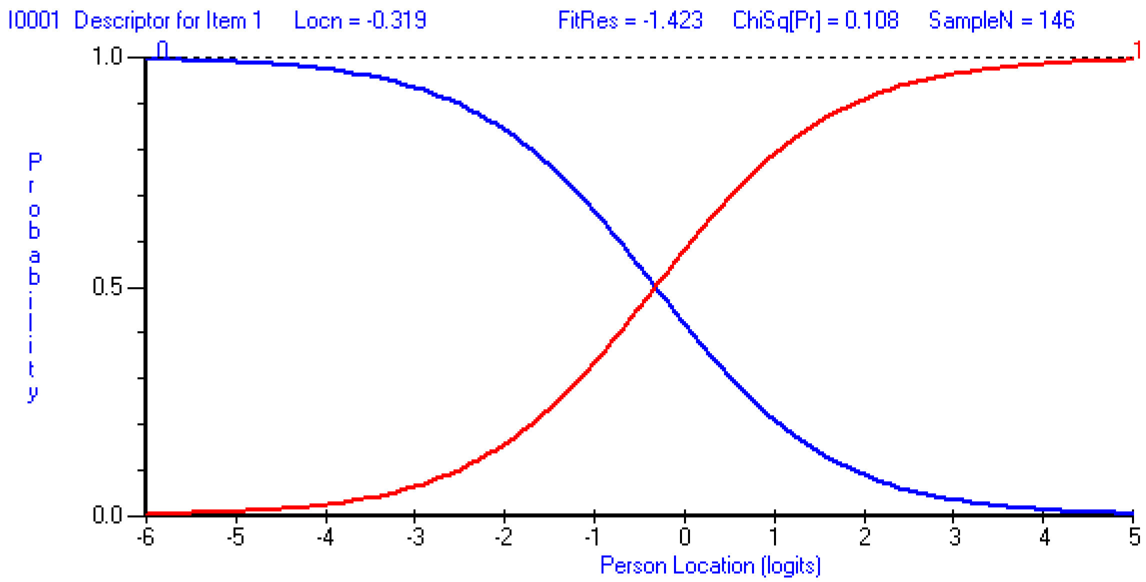


Figure 4: Response Category Curve: Item1 – Visual Discrimination Upper Case Letters

Characteristics of the Sample (VDUCL)

The measures for Visual Discrimination of Upper Case Letters were displayed in a graphical format separated by gender (Figure 5), type of school (Figure 6), age (Figure 7), grade (Figure 8) and whether intervention had been received (Figure 9). The mean differences were then tested for statistical significance using t-tests. Females have a higher mean measure than males for Visual Discrimination of Upper Case Letters but this is not statistically, significantly different (t=1.05, df=321, p=0.15). Public school students have a higher mean measure than private school students for Visual Discrimination of Upper Case Letters and this is statistically, significantly different (t=2.63, df=322, p=0.005). As would be expected, the mean measures generally increased by age from Four years of age (lowest) to nine years of age (highest) and this was statistically, significantly different (t=5.07, df=66, p<0.000). Again, as expected, the mean measures generally increased by grade from Pre-primary (lowest) to Year 3 (highest) and this was statistically, significantly different (t=8.27, df=127, p<0.000). While the mean measures for no intervention were higher than for intervention, this was not statistically, significantly different (t=1.44, df=322, p=0.07).

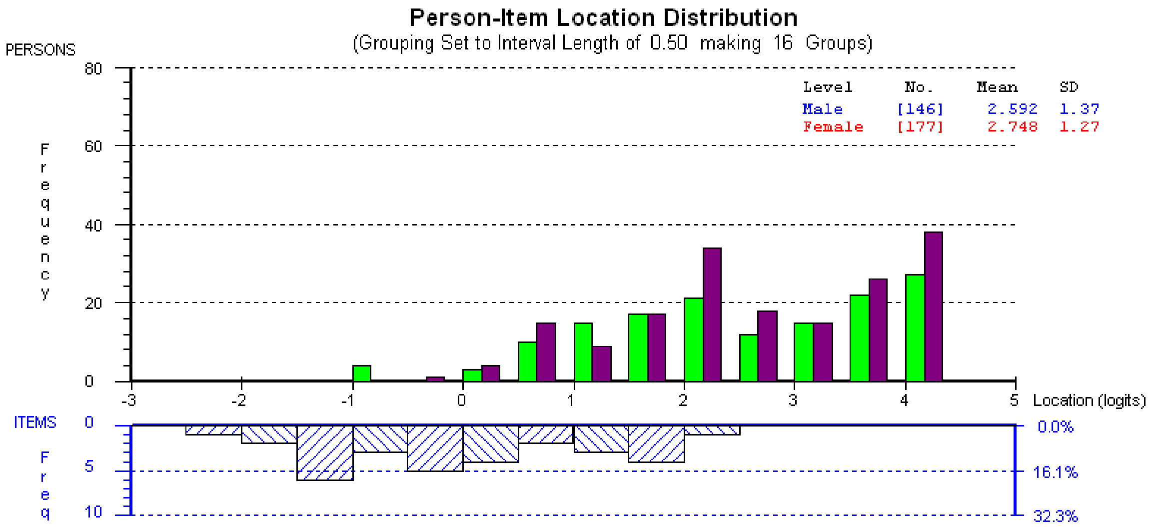


Figure 5: Target Graph by Gender for Visual Discrimination for Upper Case Letters
Note: There is a colour error in the RUMM program. Purple represents the females (not red) and green represents the males (not blue).

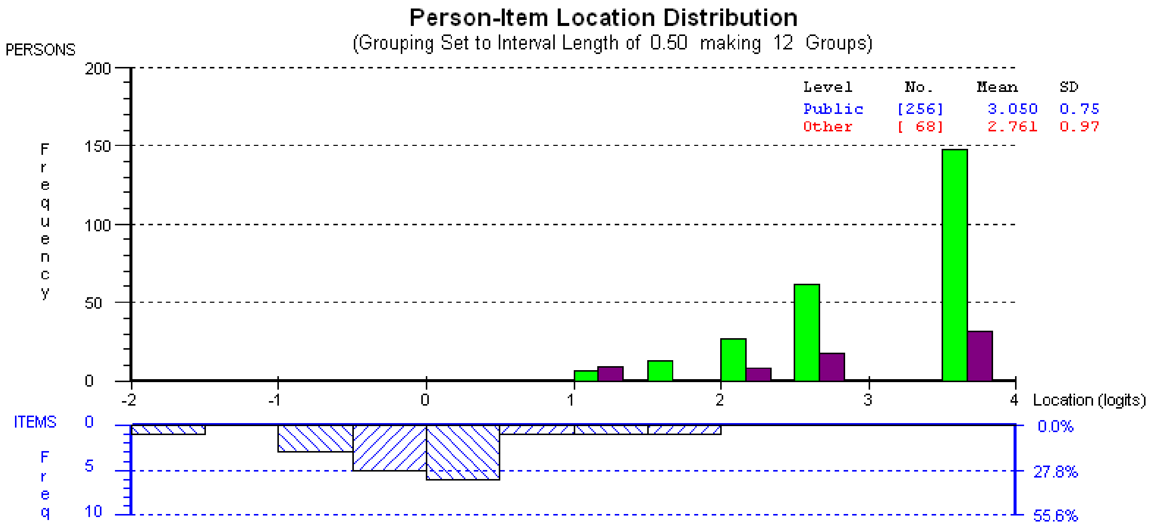


Figure 6: Target Graph by Type of School for Visual Discrimination for Upper Case Letters
Note: There is a colour error in the RUMM program. Purple represents other schools (not red) and green represents the public schools (not blue).

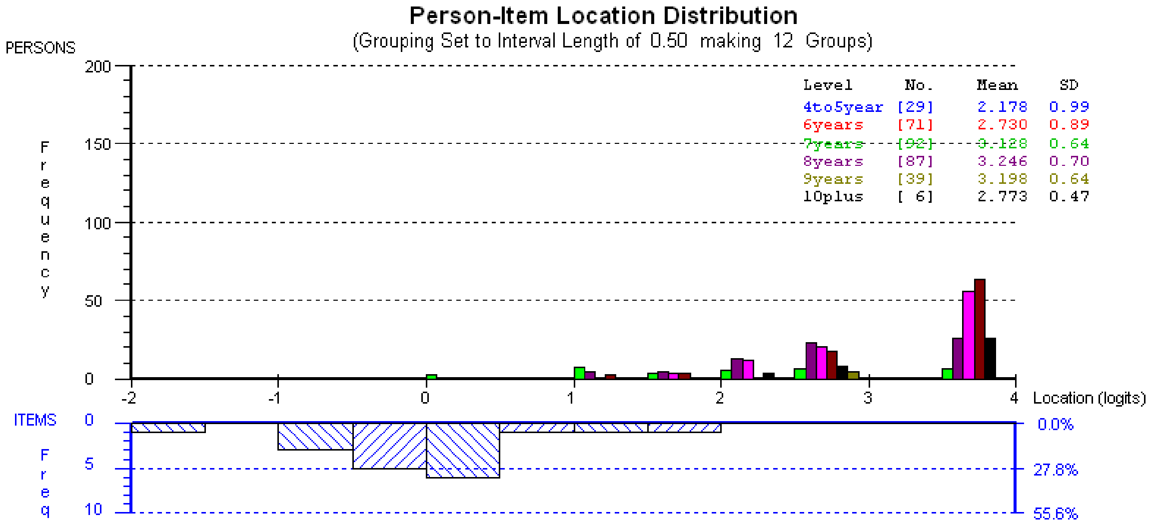


Figure 7 Target Graph by Age for Visual Discrimination for Upper Case Letters
Note: There is a colour error in the RUMM program. Four and five year olds are represented by green (not blue), six year olds are represented by Purple (not red), seven year olds are represented by pink (not green), eight year olds are represented by maroon (not purple), nine year olds are represented by black (not brown-green) and ten years and above are represented by brown-green (not black).

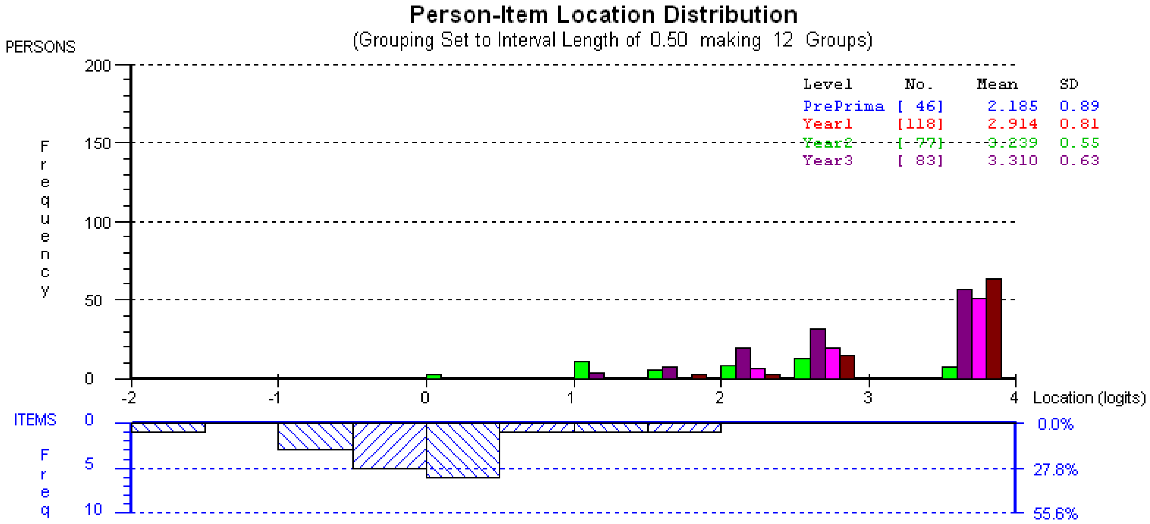


Figure 8 Target Graph by School Year for Visual Discrimination for Upper Case Letters
Note: There is a colour error in the RUMM program. Pre-primary is represented by green (not blue), Year 1 is represented by purple (not red), Year 2 is represented by pink (not green), and Year 3 is represented by maroon (not purple).

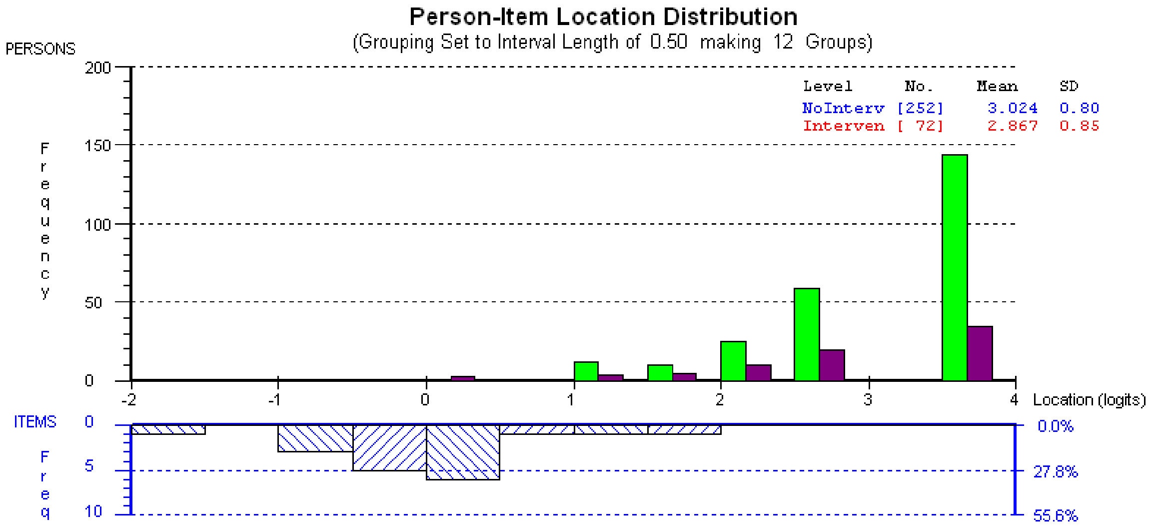


Figure 9 Target Graph by Intervention for Visual Discrimination for Upper Case Letters
Note: There is a colour error in the RUMM program. Green represents no intervention and purple intervention.

The graphical data for Visual Discrimination of Lower Case Letters was checked in the RUMM computer program but is not produced here to avoid too much repetition but the graphs are similar to those produced for Visual Discrimination of Upper Case Letters.

Females have a higher mean measure than males for Visual Discrimination of Lower Case Letters but this is not statistically, significantly different (t=1.06, df=321, p=0.15). Public school students have a higher mean measure than private school students for Visual Discrimination of Lower Case Letters and this is not statistically, significantly different (t=0.90, df=321, p=0.19). As would be expected, the mean measures generally increased by age from four years old (lowest) to ten year old or older (highest) and this was statistically, significantly different (t=10.01, df=66, p<0.000). Again, as expected, the mean measures generally increased by grade from Pre-primary (lowest) to Year 3 (highest) and this was statistically, significantly different (t=15.98, df=127, p<0.000). While the mean measure for no intervention was higher than for intervention, this was not statistically significantly different (t=1.24, df=321, p=0.10).

Discussion

The final 18 items and their difficulties are presented, in order from easiest to hardest, in Table 4 for Visual Discrimination of Upper Case Letters. The students found it easy to discriminate upper case letters that were symmetrical around the midline, for example the T, X, Y. They found it moderately easy to discriminate upper case letters that had an upright line on the left of the letter (e.g. E, R, B), moderately difficult to discriminate upper case letters that were rounded (e.g. S, G, U) and most difficult to discriminate upper case letters that were in a reversed orientation (e.g. ꟼ, ꓘ, Ǝ).

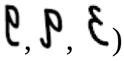
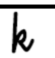
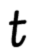
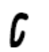
In the Visual Discrimination of Lower Case Letters (see Table 5 for the 31 item difficulties ordered from easy to hard), students found it easy to discriminate letters that began with a long downward stroke on the left, such as the k, h, b, and moderately easy to discriminate lower case letters that only consisted of a body, for example o, r, u, c. Lower case letters that consisted of only a body and were also reversed were moderately difficult to discriminate, for example ; while lower case letters with a body as well as a tail and in the reversed orientation (e.g. ) were the most difficult to discriminate.

Table 5
Difficulties for 18 Final Items in Visual Discrimination for Upper Case Letters Scale

Item No	Item Letter	Difficulty	SE	Item No	Item Letter	Difficulty	SE
2 (easiest)	T	-1.58	0.51	17	S	+0.02	0.27
24	X	-0.92	0.39	11	P	+0.11	0.27
27	Y	-0.57	0.34	15	G	+0.17	0.26
18	M	-0.56	0.34	23	U	+0.31	0.25
25	Q	-0.43	0.32	4	L	+0.32	0.25
5	A	-0.35	0.31	3	ꟼ	+0.39	0.24
1	E	-0.32	0.31	9	D	+0.66	0.22
22	R	-0.18	0.29	8	ꓘ	+1.29	0.19
28	B	-0.18	0.29	30	Ǝ (hardest)	+1.79	0.18

Note: Items are ordered from easiest (item 2, -1.58 logits) to hardest (item 30, +1.79 logits)

Table 6
Difficulties for 31 Final Items in Visual Discrimination for Lower Case Letters Scale

Item No	Item Letter	Difficulty	SE	Item No	Item Letter	Difficulty	SE
9 (easiest)		-2.06	0.42	23		-0.11	0.20
8		-1.83	0.38	30		+0.11	0.19
6		-1.51	0.33	3		+0.12	0.19
10		-1.30	0.31	2		+0.26	0.18
29		-1.20	0.29	26		+0.34	0.18
1		-1.15	0.29	7		+0.81	0.16
13		-1.10	0.28	14		+0.99	0.16
16		-1.08	028	15		+1.12	0.15
19		-1.06	0.28	33		+1.28	0.15
34		-0.90	0.26	12		+1.46	0.15
31		-0.70	0.25	24		+1.59	0.14
28		-0.54	0.23	32		+1.62	0.14
17		-0.49	0.23	36		+1.78	0.14
21		-0.35	0.22	11		+1.93	0.14
35		-0.21	0.21	4(hardest)		+2.26	0.14
25		-0.11	0.20				

Note: Items are ordered from easiest (item 9, -2.06 logits) to hardest (item 4, +2.26 logits)

Discussion on the Non-Fitting Items

Eighteen items were deleted from the Visual Discrimination Upper Case Letters due to poor fit to the Rasch measurement model. Usually the main reason for non-fit is poor agreement in regard to the item difficulty for students who have similar measures. For example, half of the medium ability students say an item is easy and half say that it is hard, thus it does not fit the measurement model and is deleted. The 12 items deleted in Visual Discrimination of Upper Case Letters were: J, H, and the reversed letters C, B, F, S, R, Z, L, N, J, and D. The students disagreed on the difficulty of these letters but the reasons for the disagreements are unknown. It is also of particular interest that most of the letters deleted due to disagreement were the letters printed in the reversed orientation. A number of students suggested that they confused the reversed J and L with the correctly oriented letter L and J respectfully because they have difficulty remembering 'which is which'. A substantial number of students requested information assisting with identification of the reversed letter J, asking "what letter is this".

In Visual Discrimination of Lower Case Letters, five of the original 36 letters were deleted due to non-fit to the Rasch measurement model. The deleted letters were the reversed letters y, j, r, f, and b. It is again noticeable that all the letters where there was poor fit were the reversed letters. Except for the letter r, the font should not have affected the students' interpretation of these letters; however, the orientation of the letters may have been a small confusing factor but it is more likely that there was another unknown reason.

Inferences from the Measures

Linear scales were created that show good fits to the measurement model and valid inferences can now be made about the students who need help. The bottom 19 student measures for Visual Discrimination of Upper Case Letters have been taken because these students all scored 14/18 or less, meaning that they were the students who responded incorrectly to the last four letters including the reversed letters. These student measures are presented in Table 7 and identified only by number for ethical reasons.

Table 7: Lowest 19 Student Measures Visual Discrimination Upper Case Letters

ID	Raw score	Location	SE	Residual
74	4	-1.32	0.58	2.31
37	7	-0.49	0.51	2.64
323	10	0.23	0.50	1.15
64	10	0.23	0.50	0.21
80	13	1.00	0.55	1.70
76	14	1.31	0.59	-0.84
74	14	1.31	0.59	1.34
42	14	1.31	0.59	-0.28
164	14	1.31	0.59	-0.02
324	14	1.31	0.59	0.29
62	14	1.31	0.59	-0.61
83	14	1.31	0.59	-0.55
27	14	1.31	0.59	0.72
72	14	1.31	0.59	-0.73
79	14	1.31	0.59	0.63
81	14	1.31	0.59	0.17
66	14	1.31	0.59	-0.64
209	14	1.31	0.59	0.20
5	14	1.31	0.59	-1.14

The child who scored four in Visual Discrimination of Upper Case Letters was only able to discriminate letters that were symmetrical around the vertical axis. Students who scored 10 had some difficulty discriminating asymmetrical letters as well as reversed upper case letters, whereas the students who scored 14 mainly found the reversed letters difficult to discriminate. Students scoring poorly in Visual Discrimination of Upper Case Letters have difficulty discriminating when upper case letters are reversed and may need extra assistance to improve this skill. Having made these inferences, teachers could tailor-make remedial work for the identified students.

The bottom 21 student measures for Visual Discrimination of Lower Case Letters have been taken because these students scored less than 19 out of 31, meaning that they were unable to discriminate the reversed lower case letters. These student measures are presented in Table 8. Students, who scored 10, were only able to correctly discriminate the easiest 10 items in the scale and had difficulty discriminating most of the lower case letters with only a body such as the c, a, r as well as the letters with a body and tail such as g, y, p. They were unable to discriminate a lower case letter when it was in the reversed orientation. The students scoring 17 correct had difficulty with the q, s, j and all the letters presented in the reversed orientation. These student measures identify students who may require assistance to improve their skill in discrimination of the lower case reversed letters. They may also be the students who reverse their letters in reading, spelling and or writing.

Table 8: Lowest Student Measures Visual Discrimination Lower Case Letters

ID	Raw score	Location	SE	Residual
323	10	-0.97	0.43	4.05
203	10	-0.97	0.43	4.28
75	10	-0.97	0.43	4.28
164	11	-0.79	0.42	0.03
64	14	-0.28	0.41	3.89
200	17	0.22	0.42	1.33
205	17	0.22	0.42	-0.68
20	18	0.39	0.42	0.81
2	18	0.39	0.42	1.20
308	18	0.39	0.42	0.14
113	18	0.39	0.42	0.13
37	18	0.39	0.42	3.73
80	19	0.57	0.43	-0.26
82	19	0.57	0.43	-1.60
110	19	0.57	0.43	0.97
81	19	0.57	0.43	-2022
208	19	0.57	0.43	-0.50
209	19	0.57	0.43	1.17
307	19	0.57	0.43	0.34
83	19	0.57	0.43	-1.56
26	19	0.57	0.43	-0.54

These results give some direction to future research along similar lines including Rasch linear measures related to letter discrimination in sequences, in words and in phrases. It also suggests that students could be interviewed individually to try to discover why they disagree about the difficulties of the deleted items. While young students may not be able state why they cannot discriminate certain letters (it may be due to a brain or memory disfunction that students cannot articulate), some may be able to provide a clue to the answer.

References

Andrich, D. (1985). A latent trait model for items with response dependancies: Implications for test construction and analysis. In S. Embretson, E (Ed.), *Test design: Developments in psychology and psychometrics* (pp. 245-275). Orlander: Academic Press.

Andrich, D., & van Schoubroeck, L. (1989). The General Health Questionnaire: A psychometric analysis using latent trait theory. *Psychological Medicine*, *19*, 469-485.

Bailey, D. M. (1991). *Research for the health professional: A practical guide*. Philadelphia: F. A. Davis Company.

Carrow-Woolfolk, E. (1981). *CAVAT*. Hingham, MA: Teaching Resources Corporation.

Catts, H. W., & Kamhi, A. G. (1999). *Language and reading disabilities*. United States of America: A Viacom Company.

Clegg, F. (1982). *Simple statistics: A course book for the social sciences*. Cambridge, GB: Cambridge University Press.

Cooke, D. M., McKenna, K., Fleming, J., & Darnel, R. (2006). Construct and ecological validity of the Occupational Therapy Adult Screening Test (OT-APST). *Scandinavian Journal of Occupational Therapy, 13*, 49-61.

Downing, S. M. (2003). Validity: On the meaningful interpretation of assessment data. *Medical Education, 37*, 830-837.

Erhardt, R. P., & Duckman, R. H. (2005). Visual-perceptual-motor dysfunction and its effects on eye-hand coordination and skill development. In M. Gentile (Ed.), *Functional visual behaviour in children: An occupational therapy guide to evaluation and treatment options* (pp. 171-228). Bethesda, Maryland: AOTA Press.

Gardner, M. F. (1991). *Test of Pictures / Forms / Letters / Numbers / Spatial Orientation and Sequencing Skills*. Burlingame, CA: Psychological and Educational Publications, Inc.

Gardner, R. A. (1978). *Reversals Frequency test*. Santa Ana, CA: Optometric Extention Program Foundation, INC.

Goldstand, S., Koslowe, K. C., & Parush, S. (2005). Vision, visual-information processing, and academic performance among seventh-grade schoolchildren: A more significant relationship than we thought. *American Journal of Occupational Therapy, 59*(4), 377-389.

Green, C., & Chee, K. (1997). *Understanding ADHD. A parent's guide to attention deficit hyperactivity disorder in children.* London: Vermilion.

Gregg, N., & Scott, S. S. (2000). Definition and documentation: Theory, measurement and the courts. *Journal of Learning Disabilities, 33*(1), 5-12.

Jordan, B. T. (1990). *Jordan Left-Right Reversal Test*. Novato, CA: Academic Therapy publications.
Kirk, S. A., Gallagher, J. J., & Anastasiow, N. J. (2000). *Educating exceptional children*. Boston, MA: Houghton Mifflin Company.
Kulp, M. T. (1999). Relationship between visual motor integration skill and academic performance in kindergerten through third grade. *Optometry and Vision Science*, *76*(3), 159-163.
Lachmann, T., & Geyer, T. (2003). Letter reversals in dyslexia: Is the case really closed? A critical review and conclusions. *Psychology Science*, *45*, 50-72.
Loikith, C. C. (1997). Visual perception: Development, assessment and intervention. In M. Gentile (Ed.), *Functional visual behaviour: A therapist's guide to evaluation and treatment options.* (pp. 197-247). Rockville, MD: American Occupational Therapy Association Inc.
McDaniel, E. (1994). *Understanding educational measurement*. Dubuque, IA: WCB Brown & Benchmark Publishers.
Messick, S. (1995a). Standards of validity and the validity of standards in performance assessment. *Educational Measurement: Issues and Practice, Winter*, 5-8.
Messick, S. (1995b). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*(9), 741-749.
Oliver, B. R., Dale, P. S., & Plomin, R. (2007). Writing and reading skills as assessed by teachers in 7-year olds: A behavioural genetic approach. *Cognitive Development*, *22*, 77-95.
Richardson, P. K. (1996). Use of standardised tests in paediatric practice. In J. Case-Smith, A. S. Allen & P. N. Pratt (Eds.), *Occupational therapy for Children* (pp. 200-224). St Louis: Mosby-Year Book Inc.
Schneck, C. M. (1996). *Visual perception*. United States of America: Mosby Year Book Inc.
Schneck, C. M. (2005). *Visual perception*. St. Louis, MI: Elsevier Mosby.
Wolf, M. (2008). *Proust and the squid: The story and the science of the reading brain*. Crows Nest, NSW: Allen & Unwin.