

**Report on the Rasch analysis of the Australian Early
Development Instrument (AEDI) using
2004 data from 6 states**

**David Andrich and Irene Styles
School of Education
Murdoch University**

Contents

Summary	1
----------------	----------

Results for each of five scales

Physical Health and Well-being scale

Social Competence scale

Emotional Maturity scale

Language and Development scale

Communication Skills scale

Summary

Summary of the report on the psychometric properties of the Australian Early Development Instrument (EDI) using the Rasch model and data from 6 states

**David Andrich and Irene Styles
School of Education
Murdoch University**

The purposes of the analyses reported here are the following:

1. To report back on the validity and reliability of each of the five scales of the Australian Early Development Indicator (EDI) after modification as recommended from initial analysis of WA data which was conducted in 2004. Data have been collected from 6 states across Australia, although only three states have substantial numbers of data at present (May 2005). The scales are:

Physical Health and Wellbeing (PHWB)

Social Competence (SC)

Emotional Maturity (EM)

Language and Cognitive Development (LCD)

Communication and Skills (CS)

2. To comment on the implications of the distributions of scores and cut-off points for identifying children at risk, and on differential item functioning amongst States on one scale.

The results of the analyses indicate the following:

1. All five scales show satisfactory psychometric properties, with the first scale (Physical Health and Well-being) being the least cohesive in terms of underlying construct. Further analysis of this scale indicates that it consist of three underlying related sub-scales (named States, Skills and General). However, for the purposes of the AEDI, it should be retained as one scale.
2. In all scales the number of response categories is now appropriate and all are working as expected.
3. Many of the scales show “clumping” at points along their distributions of scores, that is there are points at which the frequency of persons increases suddenly. This is indicative of a number of items forming small subsets within the scale so that a person responds in a similar way to all the items in that subset. The clumping helps to identify cut-off points with greater certainty (for instance around the 10% cut-off points usually taken to identify children at risk).
4. In comparing the performance of children in different states, there appear to be small differences in mean scores for all except one scale (Language) which is noticeably different with one state not performing as well as the others. The differences found on the Language scale are specific to a small number of items within the scale and can be accounted for in terms of the different approaches to curricula taken by different states. Data from other tests such as Monitoring Standards suggests that these differences do not exist by the year 3 level. The possible existence of longer term effects cannot be ascertained at present and would need to be studied using a longitudinal research design.
5. Gender differences are statistically significant for all five scales with the girls being assessed higher than the boys in each case. This confirms the research-based findings on developmental differences between boys and girls, and, indirectly, the validity of the scales.
6. The reliabilities of all scales, which can be computed with the Rasch model analysis in the presence of incomplete data, are all high and acceptable for distinguishing amongst children. These are shown in Table

Reliabilities of the scales

Scale	Reliability
PHWB	0.856
SC	0.939

EM	0.930
LCD	0.915
CS	0.919

Physical health and well-being scale (PHWB)

There are 11 statements relating to health status and psychomotor skills in the PHWB scale (items A2 to A12).

Conclusions from the analysis:

1. **The categories for all items are operating according to expectations.**
2. **The fit of the items to the model is satisfactory, as is the Person Separation Index (index of reliability), and although it represents a very broad variable, it can still provide useful information for purposes of the EDI. Analysis indicates it consists of three sub-scales which may be named “States” (relating to basic physiological states), “Skills” (motor skills), and a third sub-scale of more general physical skills. It is recommended that, for the purposes of the AEDI, the full scale (three combined sub-scales) be used.**
3. **The cumulative distribution of scores indicates a clear cut-off point around 10%. The conclusion is that this set of items can provide valid and reliable measures for children in the type of age cohort from which these data were gathered, and a robust cut-off for children at risk.**
4. **Item locations are well-spread and, although the distribution is skewed with a preponderance of children at the upper limits of the scale, the scale can be considered to be targeted well if the target group is, particularly, children who may be borderline on performance on this variable.**

Item/person distribution

The graph of the distributions of item and person locations is presented in Figure PHWB. Note that there are a number of cases of extreme scores on this measure at the upper end of the scale.

Figure PHWB1. Distribution of all person and item locations for PHWB scale.

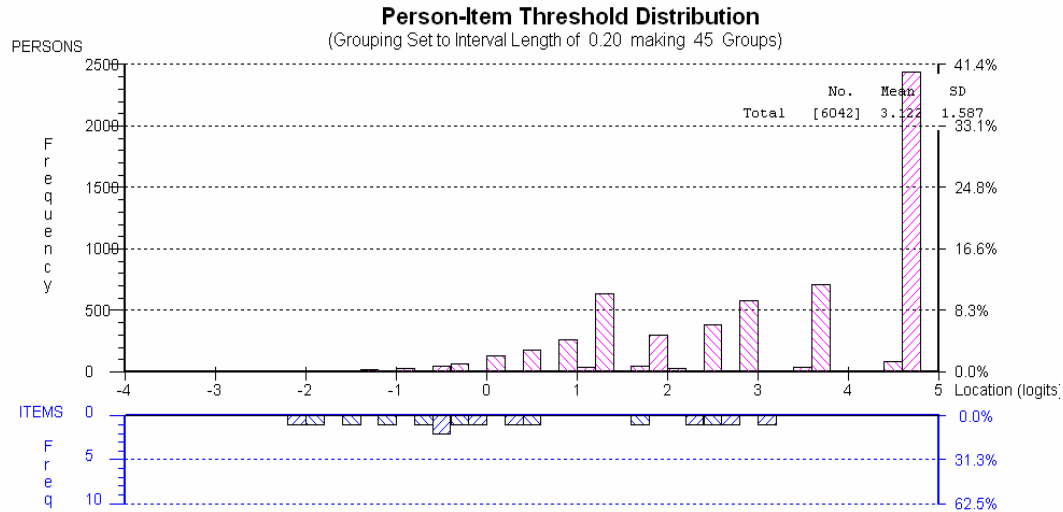


Table PHWB1. Cumulative frequency of PHWB scale scores for the combined data around the 10% cut-off point

Location	Frequency	CumFreq	CumFreq%
...
0.871	10	507	8.39
0.929	1	508	8.41
0.930	1	509	8.42
0.937	2	511	8.46
0.952	245	756	12.51
1.063	31	787	13.03
1.096	10	797	13.19
1.299	1	798	13.21
...

Table PHWB2 shows at the cut-off point 0.538 logits and 9.95% for boys, the percentage of girls at that same score is just 6.01%. In addition, it shows that if a cut off at 0.937 logits below is taken, which gives a cumulative percentage of 8.46% for the whole population, then the cumulative percentage of boys is 10.35% and girls is 6.47%. This difference is supported by the research literature on gender differences. Thus on this scale, more boys than girls would be deemed at risk.

Table PHWB2. Distribution of scores for two gender groups on the PHWB scale around the 10% cut-off point

Boys			
Location	Frequency	CumFreq	CumFreq%
...
0.322	1	207	6.84
0.416	1	208	6.88

0.431	1	209	6.91
0.538	92	301	9.95
0.585	8	309	10.21
0.861	1	310	10.25
0.871	3	313	10.35
1.299	1	798	13.21
...

Girls

Location	Frequency	CumFreq	CumFreq%
...
0.153	47	115	3.82
0.431	1	116	3.85
0.538	65	181	6.01
0.585	6	187	6.21
0.871	7	194	6.44
0.930	1	195	6.47
0.952	93	288	9.56
1.063	12	300	9.96
1.096	5	305	10.13
1.315	2	307	10.19
...

Social Competence scale

The Social Competence scale consists of 24 items (Items C1 to C23 and C25 in Section C of the instrument) which relate to a range of types of social interaction relevant to school settings.

Conclusions from the analysis:

- 1. The categories for all items worked as expected in terms of ordering.**
- 2. On the whole, the items form a scale which can provide valid and reliable measures**
- 3. The 10% cut-off score can be clearly identified.**

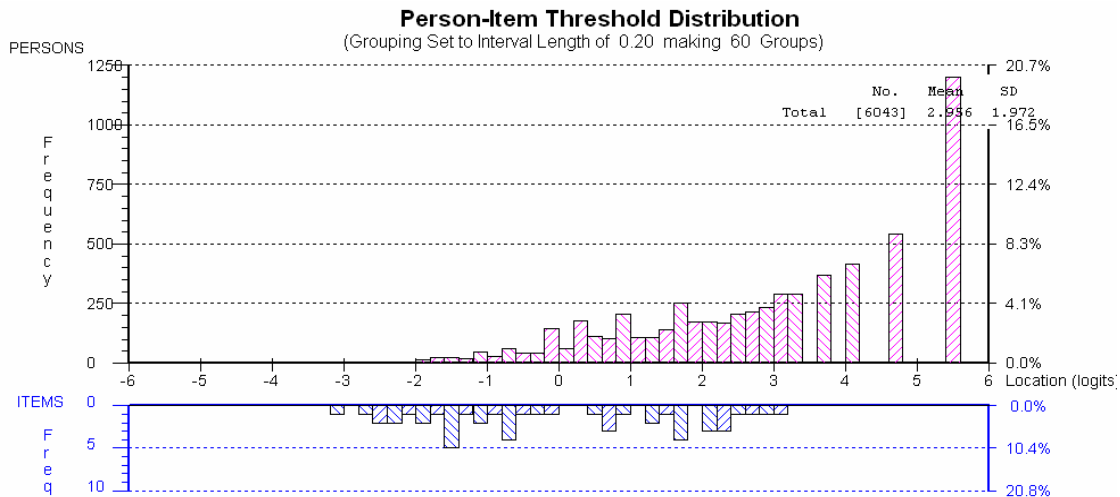
Fit of items and persons to the model

The fit of the 24 items to the Rasch model is satisfactory. The Person Separation Index was 0.939.

Item/person distribution

Figure SC1 shows the Person-item threshold distribution graph for all States. Again, there was a number of extreme scores, especially at the high-score end.

Figure SC1. Person-item threshold distribution for all States on Social Competence



The distribution graph in Figure SC1 shows that on the whole, again, items are easy relative to the persons being assessed. As discussed in the previous report, since this is an instrument that is intended to indicate readiness for school-base activities, this is not a problem, indeed it is as required because this instrument is targeted at children who are considered at a borderline acceptability and therefore at risk on this variable of social competence.

Table SC1 shows the cumulative frequency distribution for all groups in the region of the notional 10% cut-off score, with the “clumping” which is characteristic of all scales in the AEDI. Therefore it might be appropriate to take a logit score of 0.251 as the cutoff, giving 10.61% of students at risk. Again, as expected there are gender differences. Table SC2 shows these cumulative percentage differences at the cutoff logit score of 0.251. The percentage of boys below the cutoff is 14.71% and for girls it is 6.47%

Table SC1. Cumulative frequency distribution in the region of 10% cut-off score.

Location	Frequency	CumFreq	CumFreq%
...
0.211	1	552	9.13
0.235	1	553	9.15
0.245	2	555	9.18
0.251	85	640	10.59
0.301	1	641	10.61
0.397	87	728	12.05
...

Boys

Location	Frequency	CumFreq	CumFreq%
...
0.211	1	390	12.89
0.235	1	391	12.93
0.245	2	393	12.99
0.251	52	445	14.71
0.301	1	446	14.74
0.397	51	497	16.43
...

Girls

Location	Frequency	CumFreq	CumFreq%
...
0.094	2	144	4.78
0.106	18	162	5.38
0.251	33	195	6.47
0.397	36	231	7.67
0.409	1	232	7.7
...

Emotional maturity scale

The Emotional Maturity scale consists of 26 items (Items C26 to C51 in Section C of the AEDI instrument) which, as the title suggests, address the maturity and appropriateness of emotional responses in a variety of situations.

Conclusions from the analysis:

- 1. All item thresholds are working as intended with the three categories ordered clearly.**
- 2. The fit of the items to the model is good overall, as is the Person Separation Index: the conclusion is that this set of items form a single continuum and can provide valid and reliable measures.**
- 3. The targeting is similar to that for the other scales considered so far, that is, a number of respondents are beyond the level of the most difficult items and therefore would not be measured very precisely. However, again this is not a concern, given the purpose of the instrument.**
- 4. Again, the “clumping shown in the distributions makes it easy to identify the 10% cut-off point for children at risk.**

Fit of items and persons to the model

The response categories are performing according to expectations and the fit of the 26 items to the Rasch model is satisfactory. The Person Separation Index was 0.930.

Item/person distribution

Figure EM1 shows the Person-item threshold distribution graph for all States. Again, there were a number of extreme scores, especially at the upper end.

Figure EM1. Distribution of all persons and items on the EM scale.

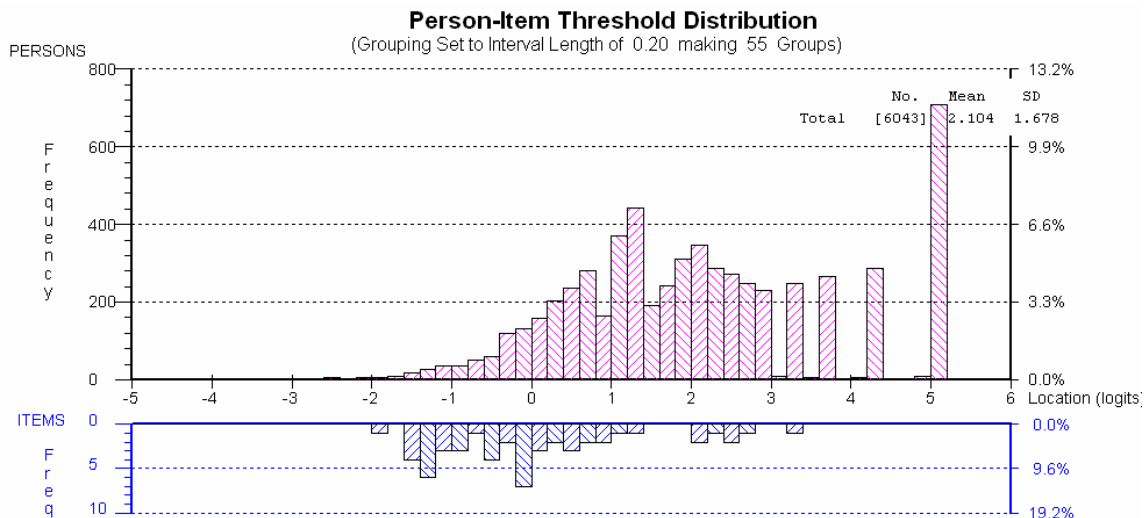


Table EM1 shows the cumulative frequency distribution for all groups around the 10% cut-off score. Taking into account the group with 70 cases, the cutoff score might be chosen at 0.162 logits with a cumulative frequency of 11.07%. Gender differences are particularly prominent in this scale. Table EM2 shows that for this same cutoff logit score, the cumulative frequency for boys is and for girls is 5.71%.

Table EM1. Cumulative frequency distribution in the region of the 10% cut-off point.

Location	Frequency	CumFreq	CumFreq%
...
0.106	1	597	9.88
0.119	1	598	9.90
0.141	1	599	9.91
0.162	70	669	11.07
0.173	1	670	11.09
0.193	1	671	11.10
0.197	1	672	11.12
...

Table EM2. Cumulative frequency distribution in the region of the 10% cut-off point for boys and girls.

Boys

Location	Frequency	CumFreq	CumFreq%
...
0.105	1	443	14.64
0.119	1	444	14.68
0.141	1	445	14.71
0.162	52	497	16.43
0.173	1	498	16.46
0.193	1	499	16.5
...

Girls

Location	Frequency	CumFreq	CumFreq%
...
0.1	1	153	5.08
0.106	1	154	5.11
0.162	18	172	5.71
0.213	1	173	5.74
0.216	1	174	5.78
0.234	1	175	5.81
...

Language and Cognitive Development scale

The Language and Cognitive Development scale has 26 items (Items B8 to B33 of Section B of the AEDI) relating to various competencies in dealing with language such as ability to identify letters of the alphabet and writing simple words.

Conclusions from the analysis:

1. **Categories are working well.**
2. **The fit of the items to the model is good overall, as is the Person Separation Index: the conclusion is that this set of items form a single continuum and can provide valid and reliable measures.**
3. **Item locations provide a very wide range on the continuum**
4. **The differences in mean scores (and in 10% cut-off scores) amongst States is accounted for in terms of the acknowledged differences in curricula. The recommendation is for at least Items 15,16,17 and 23 not to be used for identifying children at risk in State 6 since these items identify the most marked differences in pre-school curricula amongst States, rather than areas of long-term difficulty for children in State 6. Other evidence indicates these differences no longer exist by Year 3. Longer-term impacts cannot be assessed at this stage since longitudinal studies are required to establish whether there might be such an impact. Further research on this scale, its subdivision into possible subscales, and understanding of the impact of teaching, would be instructive.**

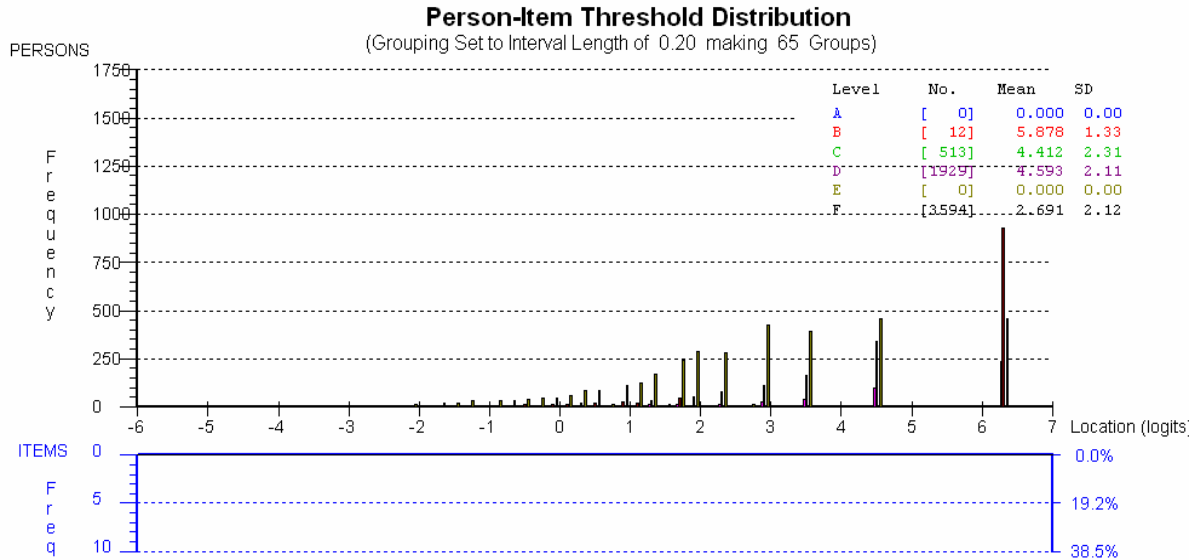
Fit of items and persons to the model

The fit of the 26 items to the model is satisfactory. The Person Separation Index was 0.915. These statistics are very good and indicate the scale fits the model and provides both valid and reliable measures.

Item/person distribution

Figure LCD1 has the graph showing distributions of item thresholds and persons by State. The mean score for State 6 is lower than those for other states.

Figure LCD1. Distribution of items and persons on Language scale by State.



The statistically significant difference in mean scores between girls and boys ($F = 140.07$, (1), $p < 0.0001$), with the girls being assessed on average higher than boys on this variable, is as expected. (Mean scores of 3.801 and 3.102 for boys and girls, respectively).

Figure LCD2. Person-item threshold distribution by Gender

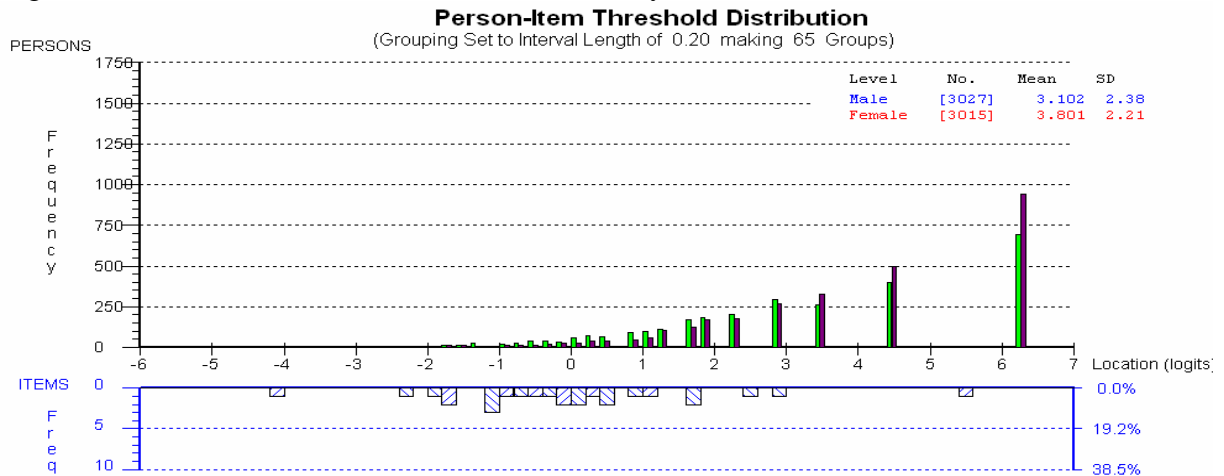


Table LDC1 shows the cumulative frequencies near the 10% cut-off point for all states. The relevant logit score of 0.260 gives a cumulative frequency of 8.33%.

Table LDC1: the cumulative frequencies near the 10% cut-off point for all states.

All

States

Location	Frequency	CumFreq	CumFreq%
...
0.204	1	499	8.25
0.230	1	500	8.27
0.238	1	501	8.28
0.238	2	503	8.32
0.260	1	504	8.33
0.334	102	606	10.02
0.377	1	607	10.04
0.439	1	608	10.05
0.445	3	611	10.10
...

Table LDC2a shows the cumulative frequencies near the logit score of 0.260 (cut-off point of all states) for 3 states with substantial numbers. It indicates that about 7.5% of State 3 children, about 5% of State 4 children, and 13% of State 6 children would be identified as being at risk. Considering the distributions somewhat further along the scale shows even greater differences which do *not* favour State 4.

Table LDC2b shows the cumulative frequencies near the 10% cut-off point for State 3. A logit score of 0.808 for State 3 gives a cumulative frequency of 9.55%. The same logit score for State 4 gives a cumulative frequency of 7.26% and for State 6 a cumulative frequency of 18.53%. This is a large difference and is explored further briefly, both substantively and statistically.

Table LCD2a. Cumulative frequency distributions for three States at a similar cut-off score corresponding to 10% for all states.

State 3

Location	Frequency	CumFreq	CumFreq%
...
0.056	1	27	5.26
0.110	8	35	6.82
0.334	4	39	7.60
0.566	4	43	8.38
0.772	1	44	8.58
0.808	5	49	9.55
1.065	8	57	11.11
1.068	1	58	11.31
1.287	1	59	11.50
1.321	1	60	11.70
...

State 4

Location	Frequency	CumFreq	CumFreq%
...
0.003	1	67	3.47
0.110	13	80	4.15
0.238	1	81	4.20
0.334	16	97	5.03
0.377	1	98	5.08
0.474	1	99	5.13
0.566	15	114	5.91
0.573	1	115	5.96
0.717	1	116	6.01
...

State6

Location	Frequency	CumFreq	CumFreq%
...
0.204	1	384	10.68
0.230	1	385	10.71
0.238	1	386	10.74
0.238	1	387	10.77
0.260	1	388	10.80
0.334	82	470	13.08
0.439	1	471	13.11
0.445	3	474	13.19
0.472	1	475	13.22

Table LDC2b: Cumulative frequency distributions for three States at a similar cut-off score corresponding to approximately 10% for State 3

State 3			
Location	Frequency	CumFreq	CumFreq%
0.566	4	43	8.38
0.772	1	44	8.58
0.808	5	49	9.55
1.065	8	57	11.11
1.068	1	58	11.31
1.287	1	59	11.5
1.321	1	60	11.7

State 4

Location	Frequency	CumFreq	CumFreq%
0.566	15	114	5.91
0.573	1	115	5.96
0.717	1	116	6.01
0.808	24	140	7.26
0.999	1	141	7.31
1.065	21	162	8.4
1.285	1	163	8.45
1.343	35	198	10.26

State 6

Location	Frequency	CumFreq	CumFreq%
0.756	1	566	15.75
0.794	1	567	15.78
0.799	1	568	15.8
0.808	98	666	18.53
0.81	1	667	18.56
0.815	1	668	18.59
0.84	1	669	18.61

Figures LCD2 and LCD3 show examples of individual items showing Differential Item Functioning against State 6– notably items 15,16,17 and23. It is recommended that, for the purposes of identifying children at risk, these items be omitted for State 6.

Figure LCD 2. Differential item functioning amongst States: Item 15: *Is able to read simple words*

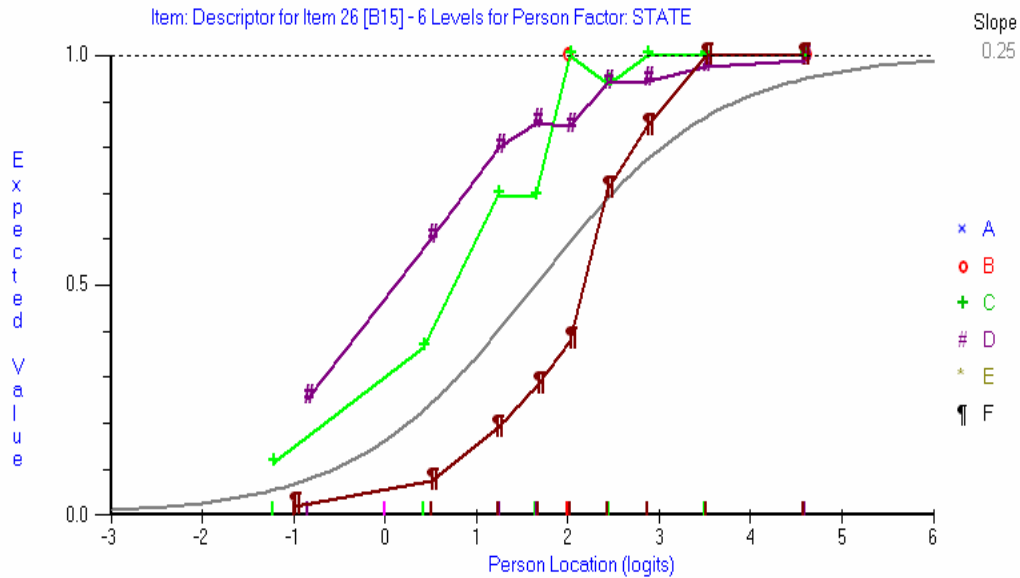
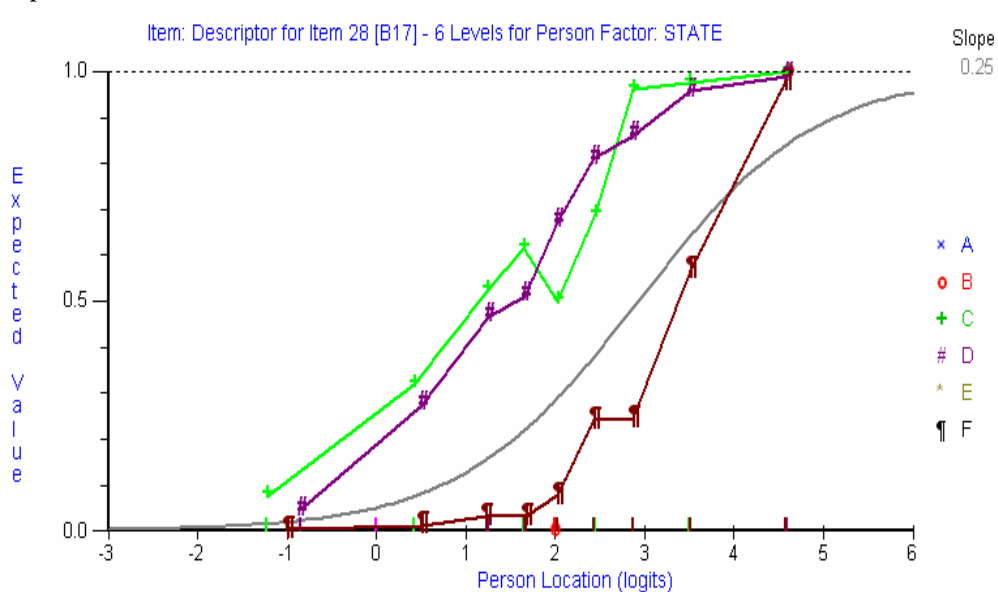


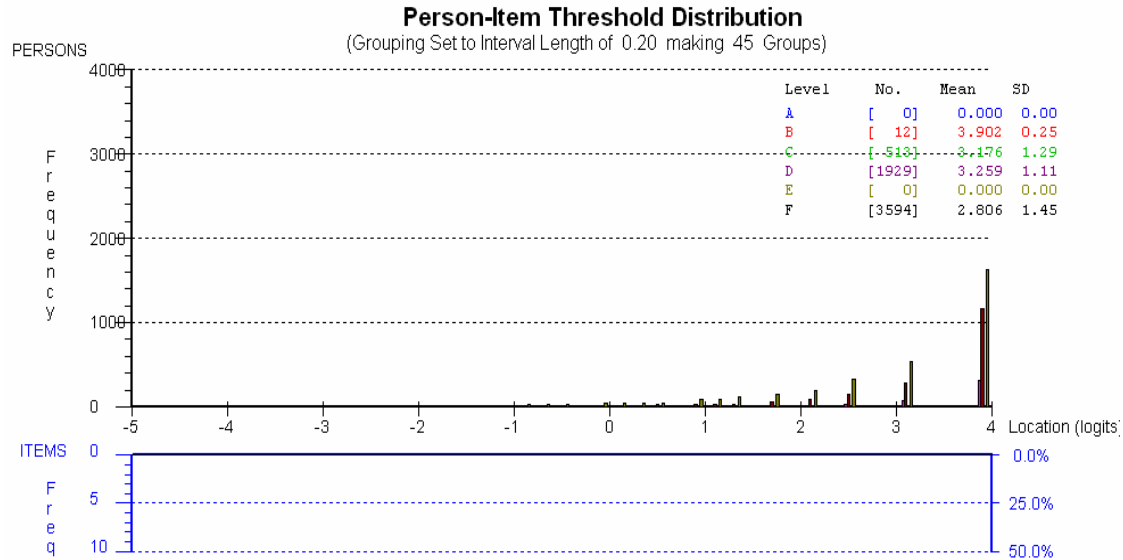
Figure LCD 3. Differential Item Functioning amongst States: Item 17: *Is able to read simple sentences*



The DIF shown in the above Figures is conditional, in the sense that given persons have the same location as evidenced by the whole scale, then the persons score differently depending in their State. Essentially, it shows that given the same total score, there is a difference by states. Because of this constraint of a total score, some other items have to show the opposite effect to some degree, either many items a small amount or a few items the opposite of that shown above. Investigations along these lines would be carried out if this scale were to be studied further in terms of its relationship to direct teaching of the material.

The person-item distribution after omitting these items, is shown in Figure LCD4. As expected, the means amongst States now show no significant differences.

Figure LCD 4. Distribution of persons and items after DIF items against State 6 are omitted.



Cumulative frequencies with the items showing DIF against state omitted are shown in Tables LCD 3 and LCD 4. The relevant logit score for the whole sample in Table LCD 3 is 0.942, with a cumulative frequency of 9.11%. For this score, the relevant state percentages now are now 7.99%, 6.58% and 11.32%.

Table LCD 3 Cumulative frequencies for all states with DIF items against State 6 omitted

All States

Location	Frequency	CumFreq	CumFreq%
...
0.891	2	546	9.03
0.893	1	547	9.04
0.933	1	548	9.06
0.935	1	549	9.08
0.942	2	551	9.11
0.993	117	668	11.04
1.075	1	669	11.06
1.098	1	670	11.08
1.109	1	671	11.09
1.116	2	673	11.13

Table LCD 4 Cumulative frequencies for 3 states with the DIF items against State 6 omitted

State 3

Location	Frequency	CumFreq	CumFreq%
...
0.516	5	28	5.46
0.698	1	29	5.65
0.750	11	40	7.80
0.942	1	41	7.99
0.993	5	46	8.97
1.251	5	51	9.94
1.496	1	52	10.14
1.529	17	69	13.45
...

State 4

Location	Frequency	CumFreq	CumFreq%
...
0.68	1	83	4.3
0.75	19	102	5.29
0.889	1	103	5.34
0.993	24	127	6.58
1.098	1	128	6.64
1.251	28	156	8.09
1.448	1	157	8.14
...

State 6

Location	Frequency	CumFreq	CumFreq%
0.893	1	404	11.24
0.933	1	405	11.27
0.935	1	406	11.3
0.942	1	407	11.32
0.993	88	495	13.77
1.075	1	496	13.8
1.109	1	497	13.83
1.116	2	499	13.88
1.123	2	501	13.94
...

Although State 6 again shows a greater cumulative frequency below the cut off point, the differences are reduced somewhat by just eliminating the 4 items. Again, further research is required with this scale.

There are a few items which show DIF between the two gender groups: these are 9 (*interest in books*) and 20 (*interest in writing voluntarily*) (in favour of girls) and 25,26,29, 30 and 31 (*interested in mathematics, sort and classify, count to 20, recognize*

numbers, recognize which number is bigger of two) (in favour of boys). The content of these items is in line with differences which would be expected from the literature on gender differences.

It is of interest that no gender DIF statistically significant with an effective sample size of 750 is discernable for the items which show DIF by State, and which may be regarded as particularly strongly influenced by formal curricula. It may be that, at these early ages, the gender differences are primarily based on differences in interests, rather than competence and that differences in competences may emerge during the course of formal schooling. A more detailed analysis by State and gender may indicate more clearly whether this could be the case.

Communication Skills scale

The fifth and last scale, Communication Skills, deals with the ability to use language to communicate with others. It has 8 items – Items B1 to B7, and Item C24

Conclusions from the analysis:

- 1. Items categories work well in terms of ordering.**
- 2. The fit of the items to the model is satisfactory, as is the Person Separation Index.**
- 3. 10% cut-off point is clearly identifiable.**

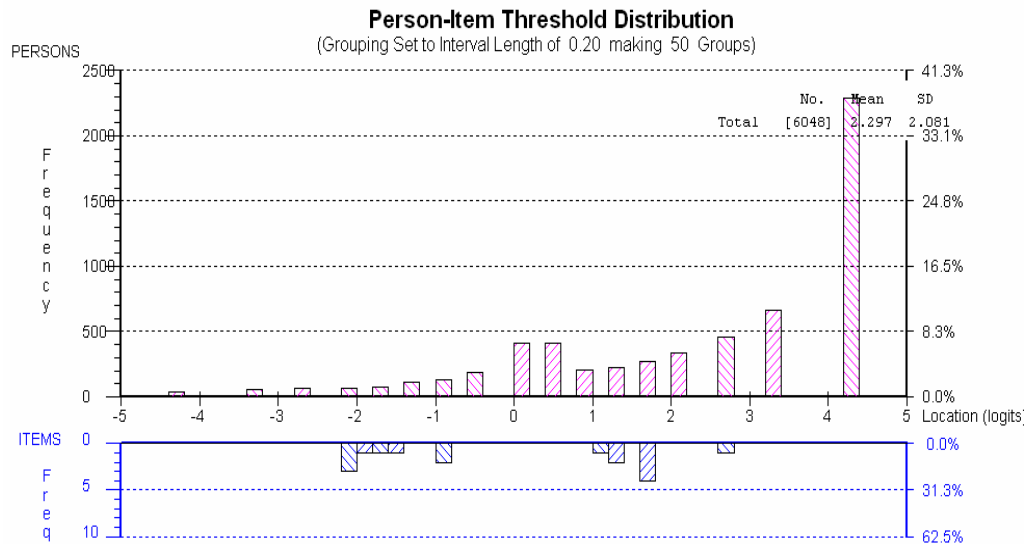
Fit of items and persons to the model

The fit of the 8 items is satisfactory as is the Person Separation Index of 0.919

Item/person distribution

The distribution of persons and items is given in Figure CS1.

Figure CS1. Person-item threshold distribution for all groups



As in the previous analysis of this scale, the distribution graph above shows that on the whole, items are targetted to the lower end of the person location distribution with a gap in the middle. This may be satisfactory if the intention is to identify only children at the lower end of the readiness continuum.

Table CS1 shows the cumulative frequency distribution in the region of the notional 10% cut-off point for this scale for all groups combined. As with all the other scale, the clumping is evident.

Table CS1. Cumulative frequency distribution in the 10% cut-off region.

Location	Frequency	CumFreq	CumFreq%
...
-0.885	128	562	9.29
-0.723	1	563	9.31
-0.626	1	564	9.33
-0.444	187	751	12.42
-0.43	1	752	12.43
-0.106	1	753	12.45
...

This scale also shows substantial gender effects. Table CS2 shows the differences in the cumulative frequencies up to approximately a logit value of -0.626 , which is the cut off which gives 9.33% cumulative frequency for the both genders. This approximate cut off for boys gives a cumulative frequency of 12.26% and for girls a cumulative frequency of only 6.37%.

Table CS2 Cumulative frequency by gender

Males

Location	Frequency	CumFreq	CumFreq%
-0.93	1	287	9.48
-0.885	82	369	12.19
-0.723	1	370	12.22
-0.626	1	371	12.26
-0.444	111	482	15.92
-0.43	1	483	15.96
0.008	245	728	24.05
0.091	2	730	24.12

Females

Location	Frequency	CumFreq	CumFreq%
-1.453	1	111	3.68
-1.413	1	112	3.71
-1.312	34	146	4.84
-0.885	46	192	6.37
-0.444	76	268	8.89
-0.106	1	269	8.92
0.008	161	430	14.26

To confirm that the Language and Development Scale was aberrant in showing a State effect, Table CS2 shows the cumulative percentages for States 3, 4 and 6 around the cut off of -0.626 logits. It is evident that the respective percentages below this cut off are very similar - 8.97%; 8.55% and 9.82% and substantially more similar than the Language and Development Scale.

Table CS2 Cumulative frequencies for States 3,4 and 6 around the same cut of score of -0.626 logits which has approximately a cumulative frequency of 9.33% below it.

State 3

Location	Frequency	CumFreq	CumFreq%
...
-1.453	1	30	5.85
-1.312	4	34	6.63
-0.885	12	46	8.97
-0.444	16	62	12.09
-0.106	1	63	12.28
0.008	49	112	21.83
0.457	39	151	29.43
...

State 4

Location	Frequency	CumFreq	CumFreq%
...
-1.737	21	88	4.56
-1.312	33	121	6.27
-0.885	43	164	8.5
-0.626	1	165	8.55
-0.444	64	229	11.87
-0.43	1	230	11.92
0.008	110	340	17.63
0.023	2	342	17.73
...

State 6

Location	Frequency	CumFreq	CumFreq%
...
-1.141	1	278	7.74
-0.93	1	279	7.76
-0.885	73	352	9.79
-0.723	1	353	9.82
-0.444	107	460	12.8
0.008	245	705	19.62

0.091	7	712	19.81
0.457	264	976	27.16
...