

TITLE: Learning Levels Of Grade V Students In Language

AUTHOR: Dr Indrani Bhaduri & Prof Avtar Singh

AFFILIATION: National Council Of Educational Research & Training, India

EMAIL: indranibhaduri@gmail.com

ABSTRACT

The National Council of Educational Research and Training (NCERT) conduct the National Achievement Surveys (NAS) at different stages of school education. The purpose of the NAS is to know the learning levels of students in subjects across the states in the country periodically and provide inputs to the policy makers, researchers and others for improving the system. Study of learning levels of grade V was initiated in 2009 and comprised of 122,543 students and 6,602 schools covering 35 States and Union Territories. Sample was drawn by using Population Proportion Sampling (PPS) and Simple Random Sampling (SRS) methods. Three forms of test booklets were developed using multiple choice items to collect data. The present study explored the learning levels in language. Item Response Theory (IRT) was used to obtain scaled scores of students for meaningful interpretation of the students knowledge and ability in language. The research findings indicate that the average achievement across the states varies significantly. Similarly across the states, within each proficiency level, students' average scores were different. To reflect upon what affects the students' learning, the background information about students, teachers and schools was also collected and analysed.

KEY WORDS: Item Response Theory, Learning Levels, Achievement Survey

Introduction

In the year 2000, the programme of National Achievement Survey (NAS), originally conceived by NCERT as an independent project, was incorporated into the Government's flagship project *Sarva Shiksha Abhiyan* (SSA). NCERT was responsible for developing and conducting the surveys whilst funding was provided by the Ministry of Human Resource Development (MHRD).

Within SSA, three cycles of NAS were planned. Each cycle was to cover three key grades i.e. Class III, Class V and Class VII/VIII. The first cycle, conducted in the period 2001-2004 was named the Baseline Achievement Survey (BAS). The second cycle, conducted during the period 2005-2008 was named the Mid-term Achievement Survey (MAS). The third planned cycle was originally named the Terminal Achievement Survey (TAS). However, the importance of these surveys and the experience gained through the first two cycles made it clear that this programme should be an on-going feature of the national education system. Therefore, the current NAS is more correctly called as the 'Cycle 3' as given in the Table 1.1 below:

Table 1: Timeline for NAS under SSA

2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Cycle 1 (formerly BAS)				Cycle 2 (formerly MAS)				Cycle 3 (NAS)			
Class V		Class III		Class V		Class III		Class V		Class III	
	Class VII & VIII					Class VII & VIII			Class VIII		

At the Class III level, two subjects are tested, namely: Mathematics and Language. For Class V, three subjects are tested: Mathematics, Language and Environmental Studies (EVS). For Class VII/VIII, four subjects are tested: Mathematics, Language, Science, and Social Science. These surveys provides a comprehensive coverage of the progress made by the education systems of the various states and union territories (UTs) of the country and also assist in planning further enhancement in the system. A total of 122,543 students from 6,602 schools in 31 states and UTs were tested in Mathematics, Language (Reading Comprehension) and Environmental Studies (EVS). The survey was aimed at evaluating the effectiveness of the schools under the SSA funding and hence only the schools financed by the state/ local governments were considered for the study. This paper reports the finding of the achievements of the student in language.

The Language tests used included three categories of item, namely:

- 'reading comprehension';
- 'language-specific elements' such as vocabulary and grammar; and
- 'extended writing' tasks.

The tests were administered in 15 languages, depending upon the students' vernacular or choice. Within these, the 'reading comprehension passages and their associated items' were translated directly and, hence, were comparable while the 'language-specific items' were, of necessity, unique to each language. The writing tasks were common across languages, however, the variations of marking was too great to allow meaningful comparisons. Therefore, this study focuses on student achievement in the domain of reading comprehension only.

Research Questions

The study aimed at finding the answers to the following research questions in order to find achievement levels in language amongst the fifth grade students:

- Are there any gender-related differences in Reading Comprehension?
- Are the differences in Reading Comprehension achievement related to the locations of the schools?
- Are there any differences in Reading Comprehension achievement related to caste (social group) category?

Methods

Setting and participants

In this research, learning levels in language, of grade fifth students was studied using a dataset that comprised of 122,543 students in 6,602 schools across 31 States and Union Territories (UT). The sample design for each state/UT involved a three-stage cluster design, which used a combination of two probability sampling methods.

- At the first stage, districts were selected using *Probability Proportional to Size (PPS)* sampling principles. This means that the probability of selecting a particular district within a state depended on the number of Class V students enrolled in that district.
- At the second stage, in the chosen districts, the requisite number of schools was selected. Once again, *Probability Proportional to Size (PPS)* principles was used so that large schools had a higher probability of selection than small schools.
- At the third stage, the required number of students in each school was selected using the *Simple Random Sampling (SRS)* method. *[In schools where Class V had multiple sections, an extra stage of selection was added with one section being sampled at random i.e. using SRS.]*

Measurements

The frameworks for 'assessment' were developed in the subject. The frameworks described:

- the competencies to be covered in the tests,
- the type of items to be used,
- the number of items to be used for testing each competency,
- the structure of the test forms and
- the number of tests.

A large number of test items were prepared and translated into the fifteen regional languages necessary for testing across the different states of India. These were then piloted in each state to see how the items worked in different languages. The difficulty level and discrimination index were computed for each item. This and other evidence collected allowed suitable items to be selected for the final tests. Three test booklets were prepared for testing the learning levels of class fifth students in language. Steps were taken to ensure that the different test booklets could be linked together. This was done by including a block of common items in each booklet. These are the 'anchor items' which, through the application of Item Response Theory, allowed placing the scores from all three booklets on the same scale.

The reading comprehension instruments used in the survey comprised six passages and thirty items, of which three were common to all three booklets, (which served as ‘anchors’ so that the different test booklets could be linked together) and one was unique to each. Each of the booklets had four reading passages with five multiple-choice items on each passage. The passages were chosen to represent a range of text types including informational passages, tables, public notices, and stories.

The items were designed to test a range of relevant *cognitive processes* or ‘reading skills’. These are classified as:

- ‘locate information’,
- ‘grasp ideas and interpret’ and
- ‘infer and evaluate’.

In ‘*Locate information*’, students were expected to find and ‘*pull out*’ a specific piece of information explicitly stated in the text. ‘Locating’ requires students to focus on a specific element of the given piece. In ‘*Grasp ideas and interpret*’ students were expected to ‘*show*’ that they have understood an idea being conveyed in the text and have interpreted it correctly. The main idea of the text, as well as, the sequence of events, the relationships between the different events/ideas/characters were required to be deciphered by the student, based on which interpretation and conclusions were expected to be drawn. In ‘*Infer and evaluate*’ students were expected to ‘*display*’ understanding beyond the information and/or ideas stated explicitly in the text. They were to comprehend the passage and infer the qualities or actions of characters; identify the text's underlying theme or message and/or evaluate its title by examining the text from more than one perspective.

Data Analysis

In earlier surveys (i.e. BAS and MAS), the learning achievement data was analysed using Classical Test Theory (CTT) and average scores were reported simply as the percentage of correct answers. This approach, whilst valid, has significant limitations. In particular, the results are linked to particular tests and groups of students so it is very difficult to use multiple tests or to link results from one year to another. Therefore, in the present study, it was decided to analyse the data using *Item Response Theory (IRT)* in addition to the classical approach. This is in keeping with the best practice of major international surveys such as Programme for International Student Assessment (PISA), Progress in International Reading Literacy Study (PIRLS) and Trends in International Mathematics and Science Studies (TIMSS). The results are reported using 'scale scores' calculated using IRT. For the present study, the chosen scale is from 0 to 500. The average score for the whole population was set at 250 and the standard deviation of the scale was set at 50 for the whole population.

The advantages of using of IRT in the present survey were as follows:

- There was an increase in the measurement points without any extra burden on the student.
- The level of difficulty of the item (question) as well as the ability of the student was placed on the same continuum.
- An item map was available so as to know the range of difficulties covered by the test.
- Unlike the classical test theory, in addition to the entire test information the item-wise information was also available.

Results

The results were grouped in three categories depending on the sampling. The three categories were those where:

- Class V students were tested and where sample coverage was adequate (>80%);
- Class V students were tested but where sample coverage was inadequate (<80%)
- Class VI students were tested at the beginning of the academic session.

The item map for reading comprehension is given below. The scale score in the first column shows the level of difficulty for each item and also depicts the minimum score on the ability scale necessary for a student to have a significant chance of success on the item. The map also includes a brief description of the *cognitive process* being evaluated. (i.e. What students needed to do to answer the item correctly.) The map shows that Class V students demonstrate a wide range of ability in the domain of Reading Comprehension.

At the lower end of the achievement scale (200-240) students are able to demonstrate all three cognitive ability as long as there the tasks are simple and there is clarity of context. For example, they are able to use information from a table, to locate the time and the occurrence of a phenomenon, recognize a particular text type, and also make simple inferences about the causes of an act described in the text.

At the intermediate range of the achievement scale (241 to 275) the students are able to, in addition to that described above, determine the causes, frequency, duration and sequence of events described in a variety of texts as well as identify cause-effect relationships and make complex inferences about the qualities of characters from their actions, and identify the effect of a given activity on the characters within a text.

At the higher end of the scale (275 and above) the students in addition to above are able to identify the main theme of a given passage and evaluate the title. They can also identify the relationship between a pronoun and the relevant object/person even when this is not immediately obvious. Furthermore, the students at this level can identify the relationships between events, ideas and phenomena, and recognize the likely thoughts of characters in the text.

Table 2: Item Map in Reading Comprehension

Scale Score	Mental Processes	Question Description
313	Infer/evaluate	Use information to identify the main theme and evaluate the title
305	Grasp ideas/interpret	Identify relationship between a pronoun and the object/person
293	Grasp ideas/interpret	Use information in the text to identify the remedy of a problem
292	Grasp ideas/interpret	identify the causal relationship of two events
285	Grasp ideas/interpret	Identify relationship between an abstract idea and a concrete phenomenon
279	Grasp ideas/interpret	Grasp the relevant idea about the thinking of people from the text
279	Grasp ideas/interpret	Use information from a notice to conclude timing of an event
279	Locate	Identify the correct place name from those given in the notice
274	Locate	Use information from a table to determine the frequency of an event
265	Grasp ideas/interpret	identify relationship between an object and its characteristics
264	Grasp ideas/interpret	Determine the sequence of activities in a process
259	Grasp ideas/interpret	Use information in the story to identify the cause of an event
258	Grasp ideas/interpret	Use information in the story to grasp the problem
255	Grasp ideas/interpret	Use information in the story to determine the cause of an event
252	Grasp ideas/interpret	Use information from the notice to derive the duration of an event
250	Locate	Use information from a table to determine the frequency of an event
249	Infer/evaluate	Make complex inference about the quality of persons from their activities
247	Infer/evaluate	Use information in the text to make inference about the cause of an act
247	Infer/evaluate	Make complex inference about the effect of an activity on concerned persons
246	Grasp ideas/interpret	Use information to draw simple conclusion about the usefulness of an object
244	Grasp ideas/interpret	Use information in the text to identify the ingredients of an object
244	Locate	Use information in a table to identify the most frequent event
242	infer/evaluate	Use information in the text to infer the cause of an event
234	Locate	Recognize appropriate information about the action to be taken by a person
231	infer/evaluate	Use information in a story to make a simple inference about the cause of an act
231	Infer/evaluate	Use information from a notice to make a simple inference about the participation in a sport
228	Infer/evaluate	Use information in a story to make a simple inference about the cause of an act
217	Grasp ideas/interpret	Recognize the text type as a notice from the format and the content
213	Locate	Use information in a table to identify the time for a phenomenon
203	Locate	Use information in a table to identify phenomena occurring at a particular time

The following three figures depict the percentile scores in reading comprehension in the three groups.

Figure 1: Percentile scores in Reading Comprehension for States where Class V students were tested and population coverage was >80%

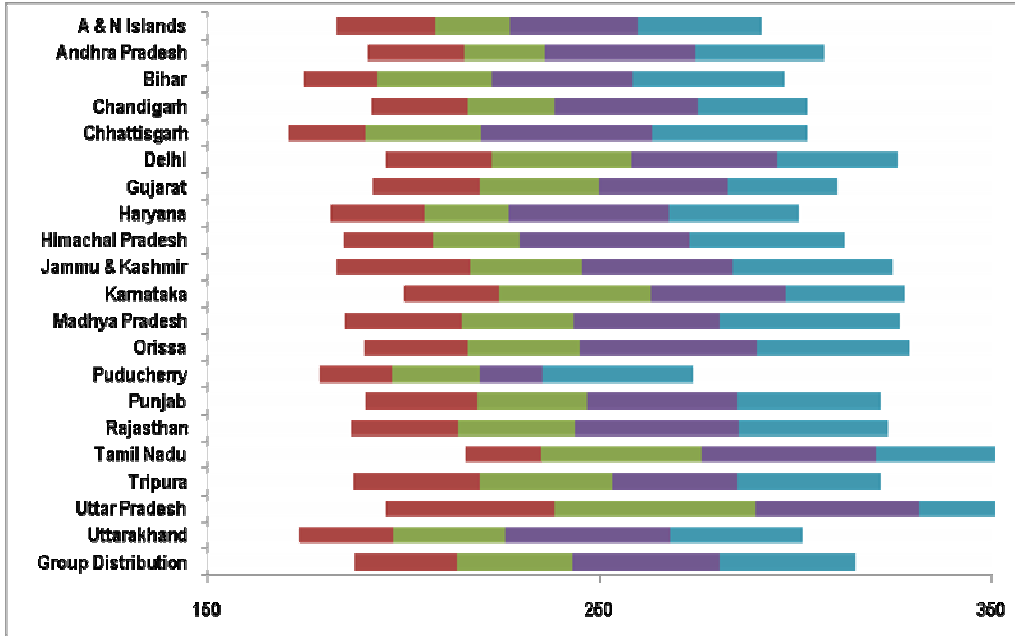


Figure 2: Percentile scores in Reading Comprehension for States where Class V students were tested and population coverage was <80%

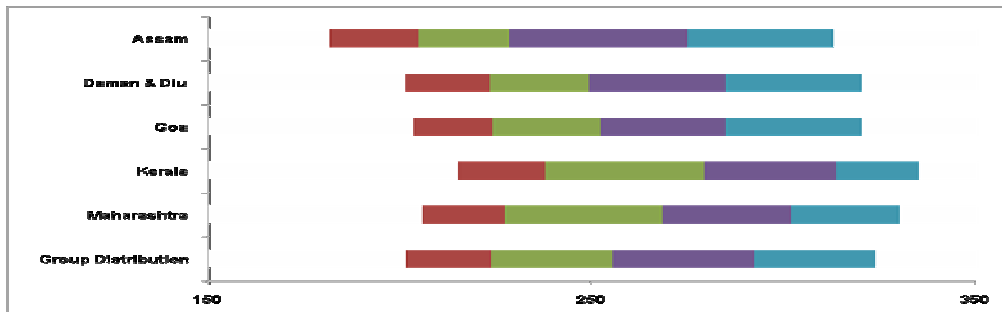
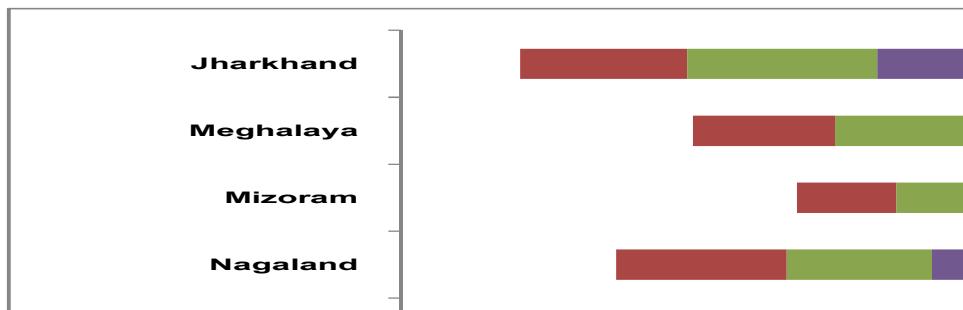


Figure 3: Percentile scores in Reading Comprehension for States where Class VI students were tested



The tables below gives the gender-wise comparison of the average 'Reading Comprehension' scores in the three groups.

Table 3: Average Reading Comprehension scores, by Gender,
for States and UTs where Class V students were tested and population coverage was >80%

State or Union Territory	Boys' Average (SE)	Girls' Average (SE)	Significant Difference
A & N Islands	229 (2.3)	237 (2.6)	♠
Andhra Pradesh	244 (2.7)	245 (2.2)	●
Bihar	229 (2.5)	226 (3.3)	●
Chandigarh	244 (2.5)	247 (2.8)	●
Chhattisgarh	229 (3.5)	228 (3.5)	●
Delhi	255 (4.1)	261 (4.8)	●
Gujarat	250 (3.0)	251 (3.4)	●
Haryana	233 (2.1)	237 (2.5)	●
Himachal Pradesh	239 (2.7)	242 (2.5)	●
Jammu & Kashmir	250 (3.0)	251 (3.1)	●
Karnataka	261 (3.1)	262 (2.9)	●
Madhya Pradesh	253 (5.0)	246 (4.1)	●
Orissa	254 (3.4)	253 (4.3)	●
Puducherry	220 (2.1)	225 (2.8)	●
Punjab	251 (3.1)	254 (2.8)	●
Rajasthan	251 (3.3)	251 (3.4)	●
Tamil Nadu	278 (2.9)	279 (2.6)	●
Tripura	252 (3.1)	253 (3.3)	●
Uttar Pradesh	284 (3.5)	283 (3.6)	●
Uttarakhand	232 (2.7)	232 (3.3)	●

Table 4: Average Reading Comprehension scores, by Gender,
for States and UTs where Class V students were tested and population coverage was <80%

State or Union Territory	Boys' Average (SE)	Girls' Average (SE)	Significant Difference
Assam	238 (2.3)	242 (2.9)	●
Daman & Diu	251 (4.4)	258 (4.1)	●
Goa	250 (4.7)	263 (5.1)	●
Kerala	272 (2.8)	280 (1.9)	♠
Maharashtra	264 (2.4)	267 (2.3)	●

Table 5 : Average Reading Comprehension scores, by gender,
for States and UTs where Class VI students were tested

State or Union Territory	Boys' Average (SE)	Girls' Average (SE)	Significant Difference
Jharkhand	239 (3.5)	234 (3.3)	●
Meghalaya	248 (2.5)	251 (2.9)	●
Mizoram	257 (1.3)	264 (1.3)	♠
Nagaland	251 (2.9)	248 (3.4)	●
Sikkim	246 (1.8)	245 (2.1)	●
West Bengal	268 (2.8)	263 (2.8)	●

- No significant difference between the average performance of girls and boys.
- ♠ Girls' average performance is significantly greater than that of boys.

Table 5 below compares the average Reading Comprehension scores achieved by students in rural and urban schools.

Table 6: Average Reading Comprehension scores by Location for States and UTs where Class V students were tested and population coverage was >80%

State or Union Territory	Rural Average (SE)	Urban Average (SE)	Significant Difference
A & N Islands	239 (3)	225 (2.7)	⤵
Andhra Pradesh	245 (2.5)	241 (3.5)	●
Bihar	227 (2.7)	238 (10.7)	●
Chandigarh	244 (3.8)	245 (3.1)	●
Chhattisgarh	228 (3.3)	238 (12.3)	●
Delhi	263 (5.7)	257 (4)	●
Gujarat	249 (2.8)	261 (8.7)	●
Haryana	234 (2.1)	243 (3.8)	●
Himachal Pradesh	241 (2.7)	240 (6)	●
Jammu & Kashmir	251 (2.7)	240 (12.6)	●
Karnataka	263 (3)	259 (6.3)	●
Madhya Pradesh	250 (4.3)	244 (8.6)	●
Orissa	251 (3.5)	280 (12)	⤵
Puducherry	220 (2.3)	227 (3.8)	●
Punjab	253 (3)	249 (5.8)	●
Rajasthan	251 (3.5)	251 (5.7)	●
Tamil Nadu	278 (2.9)	279 (5.5)	●
Tripura	252 (3.4)	255 (5.8)	●
Uttar Pradesh	283 (3.4)	278 (15.2)	●
Uttarakhand	233 (2.7)	227 (11.9)	●

Table 7 : Average Reading Comprehension scores, by Location, for States and UTs where Class V students were tested and population coverage was <80%

State or Union Territory	Rural Average (SE)	Urban Average (SE)	Significant Difference
Assam	238 (2.2)	256 (8.4)	⤵
Daman & Diu	257 (5)	245 (7.8)	●
Goa	253 (6.6)	262 (5.8)	●
Kerala	276 (2.1)	283 (4.4)	●
Maharashtra	265 (2.3)	269 (4.7)	●

Table 8: Average Reading Comprehension scores by Location for States and UTs where Class VI students were tested

State or Union Territory	Rural Average (SE)	Urban Average (SE)	Significant Difference
Jharkhand	237 (3.3)	235 (9.5)	●
Meghalaya	250 (3.3)	250 (4.1)	●
Mizoram	258 (1.6)	264 (1.6)	⤵
Nagaland	251 (3.3)	238 (6.2)	●
Sikkim	242 (1.9)	255 (2.6)	⤵
West Bengal	265 (2.6)	267 (5.4)	●

● No significant difference between the average performance of rural and urban students.

⤵ Rural students' average performance is significantly lower than that of urban students.

Table below compares the average Reading Comprehension scores achieved by students in different social categories.

Table 9: Average Reading Comprehension scores for groups by social category (Class V)

	Average (SE)	SC	ST	OBC	General
SC	245 (1.0)	-	Ⓜ	●	Ⓜ
ST	240 (1.6)	Ⓜ	-	Ⓜ	Ⓜ
OBC	247 (1.0)	●	Ⓜ	-	Ⓜ
General	253 (1.2)	Ⓜ	Ⓜ	Ⓜ	-

- The average scores of the two categories being compared are not significantly different.
- Ⓜ The average score of the category given in the first column is significantly higher than that of the category with which it is being compared.
- Ⓜ The average score of the category given in the first column is significantly lower than that of the category with which it is being compared.

Conclusion

The average achievement of students in Reading Comprehension varies greatly across the States and UTs of India. There is a highly significant difference between outcomes in high scoring States such as Uttar Pradesh (282), Tamil Nadu (278) and Kerala (277), and low scoring States/UTs such as Puducherry (222), Bihar (228) and Chhattisgarh (229). States also vary greatly in the range between their lowest and highest achieving students as revealed by their inter-quartile score ranges. Some States/UTs, e.g. Puducherry (39), Sikkim (44) and A & N Islands (51) have relatively homogeneous cohorts whilst others have far more diverse outcomes e.g. Uttar Pradesh (93), Tamil Nadu (85) and Jharkhand (81). Overall no significant differences were detected in the average achievement of girls and boys. Similarly, no significant difference was detected between the achievement level of rural and urban students although exceptions were found in a small number of States/UTs. The survey did find that students from the General Category outperformed their peers in the SC, ST and OBC categories by a statistically significant margin.

Practical Implications

This survey has implications towards assisting the educational policy framers and administrator to see that the enormous resources allocated towards the education has optimal utilisation by all the states and UTs. It also facilitate in formulating possible intervention measures that may be instituted to achieve equality and comprehensive development in the field of education throughout the nation, as envisaged in SSA. The unfortunate and persistent gender and social inequality that continue in the Indian culture and mindset must also be adequately addressed and this study is an important step towards the same.

Acknowledgments

The present study is a part of the National Achievement Survey conducted by Department of Educational Measurement and Evaluation, NCERT. We would like to thank the following people who helped with the analysis of the results and the writing of the report (in alphabetical order): George Bethell, Eugene Gonzalez, Mary Pitoniak, Mee Young, and Mamta Agrawal.

References

- Avtar Singh et al., (2006). Learning Achievement of Class V Children - A Baseline Study, NCERT.
- Avtar Singh et al., (2007). Learning Achievement of Class III Children -A Baseline Study, NCERT.
- Avtar Singh et al., (2008). Learning Achievement of Class V Children - A Mid Term Study, NCERT.
- Avtar Singh et al., (2008). Learning Achievement of Class III Children - A Mid Term Study, NCERT.
- Andrich, D. (1988). *Rasch Models for Measurement*. Newbury Park, CA, Sage.
- Beaton, A.E. (ed.) (1987). *Implementing the new design: The NAEP 1983–84 technical report*. Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.
- Foy, P. & Olson, J.F. (2009). *TIMSS 2007 user guide for the international database*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Foy, P. Galia, F & Isac, L. (2008). *Scaling the Data from the TIMSS 2007 Mathematics and Science Assessments*. Chapter 11 in Olsen et al (eds.) TIMSS 2007 Technical Report.
- Foy, P. & Olson, J. F. (2009). *TIMSS 2007 User Guide for the International Database*. Chestnut Hill, MA: Boston College.
- Hambleton, R. & Swaminathan, H (1985). *Item Response Theory: principles and applications*. Boston: Kluwer-Nijhoff.
- Foxman, D., Hutchison, D. & Bloomfield B. (1991). *The APU Experience*. HMSO.
- Leeson, H. and Fletcher, R. (n.d.) *An Investigation of Fit: Comparison of the 1-, 2-, 3-Parameter IRT Models to the Project as Ttle Data*. Unpublished working paper. Massey University, Albany Campus New Zealand.
- Olson, J. F., Martin, M. O. and Mullis, I. V. S (2009). *TIMSS 2007 Technical Report (revised 2009)*. Boston: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Thissen, D. & Wainer, H. (2001). *Test Scoring*. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- Zimowski, M., Muraki, E., Mislevy, R. & Bock, D. (1996). *BILOG-MG* (Computer program). Available from Scientific Software International, Inc. 7383 N. Lincoln Avenue, Suite 100 Lincolnwood, IL 60712-1747 USA.