EXECUTIVE SUMMARY

REPORT

THIRD INTERNATIONAL MATHEMATICS AND SCIENCE STUDY – 2003

1st NATIONAL RESEARCH CO-ORDINATORS MEETING

25-28 FEBRUARY 2001, HAMBURG, GERMANY

SJ HOWIE

TIMSS-2003
SUMMARY REPORT OF TIMSS
1ST NATIONAL RESEARCH CO-ORDINATORS MEETING
HAMBURG, GERMANY, FEBRUARY 2001

1. INTRODUCTION TO TIMSS AND ITS DESIGN

The Third International Mathematics and Science Study (TIMSS) is an international comparative study aimed at evaluating the achievement in pupils in mathematics and science across countries. TIMSS and its repeat TIMSS-R were conducted under the auspices of the International Association for the Evaluation of Educational Achievement (IEA).

The IEA Third International Mathematics and Science Study (TIMSS-95) was conducted in 1994 and 1995, and its replication TIMSS-R (TIMSS-99) followed in 1998/1999. Both TIMSS and TIMSS-R have basically the same design and data were collected through mathematics and science tests, questionnaires to schools, teachers and pupils, and a curriculum analysis project (only in TIMSS).

Three populations were tested in TIMSS-95. These are described as populations 1, 2 and 3. Population 1 comprises the students in the pair of adjacent grades that contained the most students who were 9-year-olds at the time of testing. Population 2 consists of students in the pair of adjacent grades containing the most students who were 13-years old at the time of testing. Finally, population 3 includes students in the last year of secondary school, regardless of the type of programme in which they were enrolled (Roitaille and Garden, 1996). Mathematics and science curriculum-driven achievement tests were administered to students in Population 1 and Population 2, whilst mathematics and science literacy tests were administered to students in Population 3. The TIMSS design required that a minimum of 150 schools be selected per population group. All participating TIMSS countries were required to participate in Population 2 with Populations 1 and 3 being optional.

TIMSS used a rotation-design in which for instance in population 2, the items were grouped into 26 clusters each containing 10-18 items each. All the clusters were rotated through the eight test booklets. Some items appeared in all eight books and others in half or only one book. The books were designed to be of equal difficulty and length and were answered in 90 minutes. Every student tested completed one test book and one questionnaire, which were also randomly distributed (Robitaille and Garden, 1996; 89-97). Questionnaires were developed for the school principal, the mathematics teacher, the science teacher and the student.

Data for TIMSS were collected in SA in 1995 from pupils in grade 5 and grade 6 (international grades 7 and 8) and from pupils in final grade of secondary schooling. TIMSS-R was conducted in 1998 and focused only on grade 8 pupils. The results of TIMSS have already been published (Howie, 1997;

Howie & Hughes, 1998), while the results of TIMSS-R will be published shortly (Howie, 2001).

2. OVERVIEW OF TIMSS-2003

TIMSS-2003 will comprise mathematics and science tests at grades 4 and 8 to monitor achievement during key points in the primary and middle school years. The TIMSS project provides trends in mathematics and science achievement at the 8^{th} grade by the regular assessment of these subjects (1995-1999 and now 2003). The study will collect curriculum information, linking achievement to the content and processes studied in school mathematics and science. The classroom sampling will enable matching students' achievement to the instructional practices of their teachers. The international reports for both subjects will provide a continuous scale describing what students know and can do at four international benchmarks.

3. ORGANISATION OF TIMSS-2003

The International Association for the Evaluation of Educational Achievement (IEA) based in the Netherlands has a representative council of member countries which is responsible for initiating the study. The secretariat co-ordinates the fund raising, the country participation and quality monitoring. The IEA's Data Processing Centre in Hamburg, Germany processes data from all participating countries. The International Study Centre (ISC) at Boston College in the USA, is responsible for directing the project, the management and operations internationally. Statistics Canada in Ottawa, Canada, is responsible for sampling. The Educational Testing Service in the USA conducts the psychometric scaling. The National Research Co-ordinators participate in the development of the study and are responsible for the National Implementation of the study.

3.1 Target Populations

TIMSS-2003 will focus on:
Population 1: the upper grade of the two adjacent grades with the most 9-year olds, fourth grade in most countries,
Population 2: the upper of the two adjacent grades with most 13-year-olds, eighth grade in most countries.
One of the issues during the meeting not yet decided is the inclusion of Population 3 - the grade at which most students leave school.

3.2 Framework

An assessment framework was proposed at the meeting which describes the mathematics and science from an assessment perspective which explicitly states what will be assessed and how. The important aspects of the learning context are articulated. The framework also explains the design that will be used to implement the data collection.

The framework is based on the TIMSS curriculum frameworks used in 1995 and 1999. The TIMSS curriculum model is composed of the:

Intended curriculum –
- what are students intended to learn?
- How is the education system structured to achieve these intentions?

Implemented curriculum
- What is being taught
- How is it being taught
- School, teacher and classroom factors affecting implementation

Attained curriculum
- What mathematics and science have students learned?
- What are their attitudes and perceptions about the subject?
- How are educational opportunities distributed?

It has been updated to reflect changes in curricula and pedagogy. It will also be guided by an international panel for each subject. The development is a consultative process involving the NRCs, national committees and the expert panel members.

The mathematics framework comprises the following:

Content:
- Number
- Algebra
- Measurement
- Geometry
- Data representation and analysis

Performance expectations:
- Procedural and conceptual knowledge
- Use and application
- Reasoning

The science framework comprises:

Content
- Physics
- Chemistry
- Life science
- Earth Science
- Environmental science
- Scientific Inquiry

Performance expectations
- Factual and procedural knowledge
- Conceptual understanding
- Reasoning and analysis
- Application of process skills

Test and instrument development

Originally the design proposed was :
8 booklets, which are rotated into 14 blocks and are pencil and paper tests

4 books, which contain performance assessment tasks in mathematics and science.
However, the time involved and the nature of the tasks were of concern to the NRCs who felt that the testing time of3 hours for grade 8 was too long and was reverted tot he 1999 time of 2,5 hours. A proposal will be made containing examples for the performance tasks, which will be an international option although the ISC would prefer an integrated design. Half of the items in the booklets will be mathematics and half will be science. One-third will be trend items and two-thirds will be new items.
Countries will be asked to contribute items according to the specifications.

3.3 Questionnaires
Expert panel and NRCs will identify policy issues for special attention. The development will be done by a questionnaire development group comprising the NRCs to work with the ISC. Four questionnaires will be developed:
- curriculum,
- school principal,
- mathematics teacher and science teacher
- student

3.4 Sample sizes
A minimum of 150 schools per population group in each country must be drawn to ensure 3000-4500 students per group.

3.5 Schedule for TIMSS-2003
The following is proposed as the schedule for TIMSS-2003:

2001	
Jan-Feb 2001	Draft TIMSS assessment frameworks
Feb25-28	1st NRC meeting
March-April	Revise frameworks
May2-4	2nd expert panel meeting
June 17-20	2nd NRC meeting (Montreal)
July-Aug	Prepare, review and revise final draft assessment frameworks
Sept-Oct	Publish TIMSS-2003 assessment frameworks
Nov-Dec	3rd International NRC meeting
2002	
May 2001-Feb 2002	Test and questionnaire development
March-April	Field test
Oct-Nov	Main data collection for Southern Hemisphere
2003	
March-June	Main data collection for northern hemisphere
July Dec	Data processing and review
2004	
Prepare and review international reports	

3.6 Sampling

The NRCs have to design the national sampling plans and implement these and documentation must be made on all sampling activities. Schools have to be sampled by February 2002 and sampling weights will be calculated in 2003/2004. The sampling design is a three stage stratified cluster sample. The school will be sampled using a PPS systemic sampling method and classrooms will be sampled with equal probabilities and all students will be sampled within selected classrooms. Complete national coverage should be aimed for in the national desired target population. Exclusions should be kept to a minimum, no more than 5% of the national desired target population. National samples not meeting the requirements will be footnoted in the reports. Minimum school participation rate of 85% based on sampled schools and the minimum student participation of 85% from sampled and replacement schools or the minimum overall participation rate of 75% based on sampled schools. Samples with school participation rate of less than 65% based on sampled schools will be segregated and classrooms with a student participation rate of less than 75% will be rejected.

4. GENERAL

Much discussion was held concerning the representation of the expert panels as the NRCs felt that the lack of representation of the developing nations was unacceptable. The IEA and ISC have addressed this concern after much discussion by requesting that six nominations from Latin America, Africa and the Middle East be submitted from which four will be selected (2 for maths and 2 for science). This process will be co-ordinated by Mrs Howie during up to the 13 March by which time the names from the developing countries have to be submitted by her to the IEA's Executive Director. It was proposed by the Botswanan representatives that both names for Southern Africa be nominated from South Africa and be submitted by the South African NRC.

The timing of the tests also provoked much debate as NRCs felt that the testing period was too long. This was eventually resolved by adopting the timing used in TIMSS and TIMSS-R.

NRCs felt that the coverage of the frameworks proposed by the expert panel (especially for science) was too broad. However, it was agreed that this would remain for now until the comments on the frameworks were collected from the NRCs and their within-country experts and these analysed by the ISC.

If South Africa is to contribute to the process of item development and test design as well as the questionnaire development, it will be important for the NRC to spend a considerable amount of time soliciting and collating comments from curriculum experts in mathematics and science and submitting these to the ISC. The questionnaires in particular also require considerable time to produce reliable scales of which not many were across all the questionnaires. This issue was also raised by the NRCs of the Netherlands and South Africa during the meeting and both have offered their services to the ISC to assist in the process of developing these. In particular, more questions that are relevant to the developing countries have to be developed.

5. CONCLUSION

South Africa has the opportunity to collect trend data on student, teacher and school level now that the benchmark from the 1999 data has been established. Whilst it is unfortunate that this trend cannot be reliably established from 1995, nonetheless data that will be collected in 2002 will yield information on the cohort that was tested in 1999 and provide a trend over five years. South Africa has to decide on the issue about testing population 1, the grade 4s given that this will have to be translated into the 10 other languages given the age of the pupils. Given the timelines this seems an unlikely choice, although this option exists. As the IEA has not yet decided whether or not to go ahead with population 3, grade 12s, this is a population that is of great interest to South African policymakers and therefore should be considered by the HSRC. If the HSRC decides to go ahead with this, a letter indicating that it is interested to go include this population is required by the end of March2001 to assist the IEA in deciding whether or not to include this population.

Finally, although this is a preparatory year for TIMSS, given the strict deadlines and past experience, this is an important year for a number of critical activities to take place. Contributing to the test and questionnaire design and development will consume a lot of time, if done properly and collecting the information for the sampling is critical and will make the difference between the study being considered internationally comparable or not. Our lessons from 1995 are our strongest warning about what can happen if inadequate time is spent on the vital preparatory steps. I would urge the HSRC to ensure that the international guidelines are followed and that all the timelines are kept should it decide to continue with this project. Failure to do so will eventuate in expensive and unreliable data and wasted efforts. As we have learnt from previous experience, this is not a project that can be run on a part-time or ad hoc basis. It is best managed by a project leader with some autonomy, good communication skills, is well organised, good management skills, an understanding of large-scale research, who is willing to spend a lot of overtime during the peak times and completely dedicated to collecting quality data. In addition they must be willing and able to implement additional quality assurance measures over and above the international guidelines, given the inadequacies of the education system and the lack of resources in this type of research.