TIMSS in Ontario:
Providing Information for Educational Improvement

By
Richard M. Jones

TIMSS Background

The Third International Mathematics and Science Study (TIMSS) was a complex assessment of student achievement involving over 40 countries. The study included: three separate age-groups (9-year-olds, 13-year-olds and students in their final year of high school), two subjects (mathematics and science) and both paper-and-pencil (written) and performance (practical) tests. In addition, students responded to a questionnaire that asked them about their opinions, attitudes and interests; teachers completed questionnaires related to their academic and professional preparation, the instructional approaches they used and the subject content they taught; and princ ipals completed school questionnaires that gathered information about their schools, students and teachers. These tests and surveys were administered in 1995 and the reports were published from 1996 through 1998.[1]

Most provinces took part in the Canadian study, and Ontario was one of five Canadian jurisdictions that over-sampled (involved sufficient numbers of randomly selected schools and students to allow provincial data to be generated and reported).

At the release of the first set of Ontario TIMSS results for Grade seven and eight students in 1996, the Chief Executive Officer of EQAO, acknowledged that Ontario students had not performed as well as expected and promised the Office would analyse the data to form a more accurate picture of where Ontario students were performing well and where intervention was required.

---

[1] Readers who are interested in the results of TIMSS (1995) and/or the TIMSS Repeat project (TIMSS-R, 1999) may access international, Canadian and Ontario reports, respectively at the following Websites: http://timss.bc.edu/timss1995i/TIMSSPublications.html (Boston College), http://www.curricstudies.educ.ubc.ca/wprojects/TIMSS/Reports.html (University of British Columbia) and http://www.eqao.com/eqao/home_page/nat_int/5A1e.html (Education Quality and Accountability Office).

The wealth of data made available through participation in the TIMSS assessment provided the Office with an opportunity to commission two important research projects. One project studied five topics related to issues in provincial mathematics and science education. The research report, titled *Ontario in TIMSS: Secondary Analysis and Recommendations,* was completed in 1998. Another research project studied Ontario students' errors, misconceptions and approaches in responding to TIMSS test items. A report, titled *Ontario Profiles of Performance in Mathematics and Science,* was also completed in 1998. Instructional materials, providing ideas and practical suggestions for Ontario elementary and secondary mathematics and science teachers, were developed on the basis of this research and were released one year later.

Ontario in TIMSS: Secondary Analysis and Recommendations
In 1997, EQAO commissioned a research project, through the Ontario Association of Deans of Education (OADE), to obtain information that could be used to make recommendations to improve Ontario's mathematics and science programs, particularly at the Grade 7 and 8 levels.

Several professors and graduate students/research assistants from six Ontario universities (Lakehead University, Laurentian University, Ontario Institute for Studies in Education of the University of Toronto, Queen's University, University of Ottawa and York University) were involved in the study.

Five separate topics were addressed, each focusing on the differences in student achievement between Ontario and selected high-performing jurisdictions (e.g., Singapore, Japan, South Korea, Alberta, British Columbia), as well as jurisdictions of special interest (e.g., England, United States). The research topics were as follows: curriculum, teaching practices, teacher education, achievement correlates and Francophone sudents' results.

What follows are selected highlights from the final research report. When reviewing these findings the reader should be reminded that in the intervening years since the administration of the TIMSS assessment in 1995, there have been significant curriculum, assessment and other education-related policy changes in Ontario.

Curriculum (mathematics)

- In mathematics, the Ontario curriculum did not match well with the TIMSS tests; about one-half of Ontario teaching time, in both Grade 7 and 8 mathematics, was allotted to topics not covered by TIMSS.
- About 42% of the TIMSS items were not covered by Ontario's Grade 7 curriculum, and about 17% were not covered by the Grade 8 curriculum.
- Compared to Alberta and Singapore, Ontario results showed evidence of some specific strengths in *whole numbers and decimals* and *organization and representation of data* and particular weaknesses in *variables, formula and equations* and *measurement*.
- The link between the amount of time spent teaching mathematics and student achievement was not clear from the TIMSS data.
- Although the new (1997) intended Ontario mathematics curriculum contains greater specificity of topics, there is little indication of how much time should be spent. Consequently, there may be less correspondence with the TIMSS item pool than was the case with the 1985 curriculum.

Curriculum (science)

- There were large differences in the amount of science teaching time in Ontario as compared to higher-performing jurisdictions.
- A clearly articulated elementary science curriculum was a factor that was consistent with higher levels of performance.
- There was a relatively weak link between the intended curriculum and the attained curriculum (what students knew and could apply as measured by the achievement tests) in Ontario. The intended curriculum was a relatively poor indicator of what students actually knew and could do.
- In high-performing jurisdictions, such as Singapore and Alberta, the implemented curriculum (as actually taught by classroom teachers) was more consistent with the intended curriculum than was the case in Ontario where there was much more discretion over curriculum.
- With the exception of *chemistry* in both Alberta and Ontario and *earth science* in Ontario, each of the TIMSS reporting categories (*earth science, life science, physics, chemistry,* and *environmental issues and the nature of science)* were adequately covered in provincial curricula.

Teaching practices

- In jurisdictions where teachers had less influence/choice in determining the subject-matter they taught, the TIMSS scores were relatively high. The same was true for jurisdictions where there was a clearly defined provincial/national curriculum, and where teachers were familiar with the curriculum.
- Teachers in Alberta, Singapore, and England assigned homework in science more often than did Ontario teachers.
- There appeared to be no discernible relationship between class size and student achievement on TIMSS.
- The emphasis that the teachers in some higher-performing jurisdictions placed on certain student assessment methods, appeared to differ from Ontario.
- There was no clear relationship between the amount of preparation time, outside of the formal school day, and student achievement.
- Generally, teachers in the higher-performing jurisdictions reported feeling they were more prepared to teach the science topics listed in the TIMSS questionnaire than Ontario teachers.

Teacher education

- Few mathematics or science high school courses were required for entrance into Ontario teacher education programs.
- Unless Ontario teachers specialized in mathematics or science, their programs may or may not have included courses in these subject areas.
- In many jurisdictions, including Ontario, it appeared that while a teacher may or may not have opted to take mathematics or science as part of his/her training, he/she was regarded as qualified, on graduation, to teach all subject areas within the school curriculum.
- The qualifications for teacher certification, set out by Canadian colleges of teachers and various government agencies, make general statements about the components of teacher education programs without relating them to specific subject matter.
- In Ontario, the teacher education component of pre-service training is one year compared to two years in some other provinces.
- The TIMSS data showed that Singapore teachers have more time allotted for, or choose to spend more time on, professional development.

Achievement correlates (demographic variables)

- There was little difference between the achievement of males and females in mathematics and science.

- Students born in Canada had substantially higher achievement results in mathematics and science than students born outside of Canada.
- Whether or not the student's mother or father was born in Canada had little effect on mathematics achievement, but there was a modest effect on science achievement.
- How often students spoke the language of the test at home was related to achievement in mathematics and science.
- Those students living with both parents scored highest, followed by students living with the mother only, and then students living with the father only.
- Substantial differences in student achievement were observed across levels of parental education. The higher the parental education, the higher the students' scores.
- There was a strong relationship between the frequency of skipping class and achievement. The greater the number of reported incidents of skipping, the lower the achievement.

Achievement correlates (personal activities)

- Modest levels of time spent in various daily activities such as watching television or videos, playing computer games, reading books, playing with friends, doing jobs around the house and playing sports were associated with higher performance.
- There was a strong negative relationship between achievement and the time spent weekly in extra mathematics and science lessons, but there was a modest positive relationship between achievement and the time spent participating in mathematics or science clubs.
- Student achievement was positively associated with increases in the frequency of leisure reading.

Achievement correlates (resources)

- There was a strong, positive relationship between the number of books in the home and achievement.
- There was a strong, positive relationship between student achievement and the possession or existence of a computer, study desk, calculator and dictionary in the home.

*Achievement correlates (*attitudes*)*

- There was a strong, positive relationship between achievement and whether or not mothers thought it was important to do well in mathematics and science.
- There was a strong, positive relationship between achievement and whether or not students thought it was important to do well in

mathematics and science.

- There were strong, positive relationships between achievement and students who reported they: liked mathematics or science, enjoyed learning mathematics or science, felt that mathematics or science were important in life and would like having a job involving mathematics or science.

Achievement Correlates (teaching strategies)

- Relatively higher frequencies of some teaching strategies were associated with lower achievement. These included: copying notes from the board, having frequent quizzes or tests, working on projects and using computers.
- The frequency of working on one's own from worksheets or textbooks in mathematics was associated with higher achievement, but was mildly associated with lower achievement in science.
- The frequency of calculator use in class showed little difference in mathematics achievement, but there was a substantial negative difference in science achievement.

Achievement correlates (community factors)

- There was a strong relationship between the type of community and achievement. Students in suburban and rural areas tended to perform better than those in urban areas in mathematics and science.
- There was a strong, negative relationship between the frequency of student absenteeism and achievement in mathematics and science.
- There was a strong, negative relationship between the frequency of classroom disturbances and achievement.
- School principals' reports of the following characteristics were negatively associated with student achievement: disadvantaged economic background, neither parent with greater than primary school education, single-parent families, having learning problems and having nutritional problems.

Achievement correlates (school factors)

- Higher levels of achievement in mathematics and science were associated with higher levels of parental involvement.
- Mathematics and science achievement was associated with school principals' perceptions of shortages or inadequacies of instructional materials (e.g., textbooks), calculators and library materials. The less the perception of inadequacy, the higher the achievement.
- Students of all personal backgrounds did better when in classes with higher socio-economic status (SES) students, and did poorer when in

classes with lower SES students.

Francophone students' results (family characteristics)

- In science, there was a strong relationship between language spoken in the home and achievement. The students who "always" spoke French at home had better results than those students who spoke French "sometimes" or "never" at home.
- In mathematics, language spoken in the home was a variable that had some association with achievement, but its contribution was mediated by other socio-cultural variables such as parental education.
- There was a strong relationship between parental education and student achievement in mathematics and science. As the level of schooling increased, the results of students improved. Students of parents with a university degree had the best results.
- There was a strong relationship between the number of books in the home and student achievement in mathematics and science. In particular, students who reported there were more than 100 books in the home performed significantly better than other students. Other related socio-economic indicators such as availability of computers at home and space for the student to do schoolwork at home were also associated with achievement.

Francophone students' results (student attitudes)

- The importance that students attached to success in mathematics and science was associated with achievement. Francophone students who felt it was important to succeed tended to do better.
- Among the Ontario Francophone students, there was a relationship between the students' perception of their success in mathematics and science and their level of achievement.
- There was a relationship between mothers' attitudes about the importance of mathematics and science and achievement.
- Those Ontario Francophone students who thought that mathematics and science would be useful for their futures, tended to perform better in mathematics and science.
- A large proportion of British Columbia and Ontario Francophone students (nearly 50%) found mathematics boring. Thirty-seven percent of Alberta students and 45% of Ontario Francophone students found science boring.
- Ontario Francophone students that liked science and did not find the subject boring performed better than other students.

Francophone students' results (teaching practices)

- In general, in mathematics, there were few differences between the teaching practices of Ontario Francophone teachers and those of British Columbia. The most frequent practices included explaining concepts or principles, giving homework, demonstrating how to solve mathematics problems, solving example problems tied to new content, beginning homework in class, and following or working from a textbook.
- The practices associated with higher achievement in mathematics included giving homework, beginning homework in class, doing exercises taken from a textbook, and demonstrating problem solutions.
- In general, in science, there were some important differences between Ontario Francophone and Alberta teaching practices. Most often, Ontario Francophone teachers worked with students in small groups, presented students with problems tied to concrete everyday experiences, and seldom used a textbook, even if one was available to them. On the other hand, Alberta teachers used textbooks and standardized tests and examinations with multiple-choice questions more often than Ontario Francophone teachers.
- Ontario Francophone teachers gave their students less homework than Alberta teachers, and the homework assigned did not take as long to complete. Alberta teachers exercised strict control of homework, corrected it, and discussed it in class.
- The strategies that were positively associated with performance in science included working in small groups, solving problems linked to everyday life, and explaining rules and definitions.
- The practices that were negatively associated with achievement in science included working with a textbook, using a calculator or computer, and having students check/verify their homework among themselves. The less the teacher used these strategies, the better was success on the science test.

Ideas and Practical Suggestions for Classroom Teachers

Because TIMSS used a randomly selected sample of schools and students, rather than involving every student, it was not possible to generate results for individual schools or students. Therefore, in keeping with EQAO's policy of undertaking assessment for the purposes of improvement, the Office commissioned an analysis of the TIMSS achievement data to identify items on which Ontario students performed poorly compared to Canadian and international results. Next, these items were analysed to determine the possible errors and misconceptions that may have given rise to the incorrect answers.

Design of the study

The study involved the examination of student responses to test items from TIMSS. Classical item analysis and descriptive statistics were used to examine and compare proportions of student responses to options for multiple-choice items and to proportions of codes assigned to answers for free-response items. Results from individual items and from sets of related items were used to develop profiles of responses.

Each item corresponded to a skill or concept that it was intended to examine. Therefore, in order to determine areas of strength or weakness, the focus of the analysis was on the p-value, or percent correct. Since, in the design of multiple-choice items, incorrect options were comprised of plausible errors, an analysis of responses to these options was undertaken to determine the types of errors and misconceptions demonstrated by students. In the case of free-response items, an examination of codes assigned to incorrect responses enabled this type of analysis to be undertaken.

The study examined student responses to achievement items at each of the three levels of the system: Grade 4, Grade 8, and the final year of secondary school. Results from Ontario, Ontario (French), Ontario (English), Alberta and British Columbia were used for comparative purposes. International p-values were also used as part of the comparative analysis.

Comparisons of achievement for major topics in the curriculum at each level were undertaken at the initial stage of analysis. These involved the mean percent correct for each set of items comprising a topic. In addition to calculating the mean percent correct by jurisdiction for each topic, the standard error associated with each of these statistics was calculated so that statistically significant differences could be determined. Comparisons were then undertaken with the International mean percent correct.

When comparisons relative to item performance were made, responses were classified into three categories: items with jurisdictional p-values five or more percentage points higher than the international p-value, jurisdictional items with p-values five or more percentage points lower, and items with jurisdictional p-values between these ranges. These comparisons were made between the international p-value and each of the five Canadian jurisdictions. An additional comparison, using these criteria, was made within Ontario between Ontario (French) and Ontario (English). Five percentage points was selected as the criterion measure on the assumption that differences of that magnitude have practical implications.

Once a comparative context was established, items underwent a second process of classification. Within each jurisdiction, items were sorted on the basis of the proportions of students who answered correctly. At this point, there were two levels of focus: on items with p-values greater than or equal to 0.75 and items with p-values less than 0.50. Items in the former category were considered to reflect a strong performance and those in the latter were considered to be weak. Since each item was associated with a skill or concept, corresponding lists were established in each of the two categories for each jurisdiction. In this way, profiles of performance were developed, showing areas of strength and weakness based on item-level performance.

Based on these analyses, implications for instruction and curriculum focus for Ontario were developed for each level.

TIMSS scoring techniques and criteria

As indicated earlier, the unit of analysis on which this study was largely based was the test item. The rationale for this focus included the potential value which could be gained by studying options selected by students on multiple-choice items and the answers they provided for free-response items. This information was readily available for both types of items. An analysis of responses to multiple-choice items is available through any standard item analysis software program. In the case of free-response items, this type of analysis was possible due to the special system of coding used in TIMSS.

Multiple-Choice Items

Distractors developed for multiple-choice items were designed to reflect plausible errors. In many cases, these errors reflect common student misconceptions, lack of information, or types of calculation errors. To illustrate the potential which can be gained though the analysis of responses to options, an example from the Grade 8 level follows.

O-02	If the price of a can of beans is raised from 60 cents to 75 cents, what is the percent increase in price? A. 15% B. 20% C. 25% D. 30%

To answer this question, students needed to find the amount of the increase (15 cents) and then find that value as a percent of the original price. The correct answer to this question is *C.* A list of the incorrect options, with a possible type of error corresponding to each, is shown next.

Incorrect Option	Nature of associated error associated
A. 15%	The increase in price was 15 cents. This is not the percentage increase.
B. 20%	This is the percent of increase taken on the final sale price rather than on the original price.
D. 30%	This is twice the number of cents the tin of beans increased. It is also a number which could be a reasonable-looking answer and one which is part of a pattern.

If a substantial proportion of students selected either option A or B, some insight could be gained on the type of misconception they held.

Free-response items

There were two categories of free-response items in TIMSS: short-answer and extended-response. Short-answer items consisted mostly of single-step questions in responding to which the student wrote down an answer without needing to show work; each of these questions was assigned a weighting of one. Extended-response items were multi-step questions worth two or more marks.

The free-response items in TIMSS were intended to collect information on the processes and types of thinking which students undertake in answering questions in mathematics and science. Examination of student responses to these makes it possible to explore some of the complex and multi-stage thinking which takes place when students answer questions in these subjects.

Two-digit codes were assigned to each student response, in order to collect information on the following three aspects: correctness of response, the method or approach used in the problem and the misconception or type of error that may have been demonstrated. The first digit in each code corresponded to correctness. If a response was worth 2 marks, it received a code with a first digit of 2; if it was worth 1 mark, the code would begin with a 1. Incorrect responses began with 7, and blank or crossed-out answers began with 9. The second digit corresponded to the method or approach used if it was correct, or to the misconception or type of error if it was wrong.

As an example, the coding guide for the item shown next allowed for analysis of students' conceptions of water balance and temperature regulation of the human body.

O16 Write down the reason why we get thirsty on a hot day and have to drink a lot.

Code	**Response**
Correct Response	
10	Refers to perspiration and its cooling effect and the need to replace lost water
11	Refers to perspiration and only replacement of lost water
12	Refers to perspiration only
19	Other acceptable explanation
Incorrect Response	
70	Refers to body temperature (being too hot) but does not answer why we get thirsty
71	Refers only to drying of the body
72	Refers to getting more energy by drinking more water
76	Merely repeats the information in the stem
79	Other incorrect
Non-response	
90	Crossed out/erased, illegible, or impossible to interpret
99	Blank

Teaching resource booklets

Based on the analyses of Ontario students' achievement on the TIMSS test items, a series of teaching resource booklets was developed in both English and French for each of the three TIMSS populations: Grade 4, Grade 8 and the final year of secondary school.

The purpose of these booklets was to provide Ontario teachers with suggestions and ideas on how to more effectively teach topics, related to Ontario curriculum, which students found difficult.

The booklets are divided into mathematics and science sections, and within each section, they are organized by major content areas as defined in the TIMSS tests.

Each major content area begins with a description of the types of errors and misconceptions found through the analysis of the TIMSS results. Next, the TIMSS test item is presented with the results for Ontario. In cases where a TIMSS item must remain confidential (i.e., it had not been made public), it was replaced with a similar item or was described in detail. This information is then followed by suggestions for clarifying the concept, and in some cases, examples are given of the types of questions/activities that could be used to reinforce student understanding of the concept.

Although the "target" grades for the booklets are Grades 4, 8, and end of secondary school, some of the test items and concepts relate to other grades within the Ontario curriculum. Consequently, the Appendix of each booklet provides, for each TIMSS topic reported, the TIMSS content area, mathematics/science topic and strand, as well as the grade level where it is found in the Ontario curriculum. In this way, the resource booklets are valuable to virtually all Ontario elementary and secondary teachers.[2]

Conclusion

One of the greatest criticisms of large-scale national and international assessments is that beyond the inter-jurisdictional comparative student achievement information, they have little relevancy at the local/classroom level, and they provide limited information to inform

---

[2] Readers who are interested in viewing these teaching resource booklets may access them at the EQAO Website at: http://www.eqao.com/eqao/home_page/nat_int/5A1e.html.

educational policy making. This view, however, is mistaken and unfortunate. Such assessments often generate a great amount of rich data that are often under-used. The TIMSS data have provided Ontario policymakers and educators with useful information to support educational improvements.

References

Marshall, M.A., & Taylor, A.R. (1998). *Ontario Profiles of Performance in Mathematics and Science.* A research report prepared for the Education Quality and Accountability Office, Toronto.

Orpwood, G., Nagy, P., Roulet, G., Russell, H., Bourdignon, C., Bartley, A., Puk, T., Sanderson, N., Epp, W., Epp, J., Kirby, J., Herry, Y., Lacasse, R., & Radford, L. (1998). *Ontario in TIMSS: Secondary Analysis and Recommendations.* Research conducted for the Education Quality and Accountability Office on behalf of the Ontario Association of Deans of Education, Toronto.

Author (1999). *A focus on Effective Mathematics and Science Teaching: Grade 4 Mathematics and Science,* Toronto: Education Quality and Accountability Office.

Author (1999). *A focus on Effective Mathematics and Science Teaching: Grade 8 Mathematics and Science,* Toronto: Education Quality and Accountability Office.

Author (1999). *A focus on Effective Mathematics and Science Teaching: Intermediate Level Mathematics and Science Literacy,* Toronto: Education Quality and Accountability Office.

Author (1999). *A focus on Effective Mathematics and Science Teaching: Advanced Grade 12/OAC Mathematics and Science,* Toronto: Education Quality and Accountability Office.

Author (1999). *Pour un enseignement efficace des mathématiques et des sciences: Mathématiques et sciences en 4^{e} année,* Toronto: Office de la qualité et de la responsabilité en éducation.

Author (1999). *Pour un enseignement efficace des mathématiques et des sciences: Mathématiques et sciences en 8^{e} année,* Toronto: Office de la qualité et de la responsabilité en éducation.

Author (1999). *Pour un enseignement efficace des mathématiques et des sciences: Apprentissage des mathématiques et des sciences, Secondaire, 1^e cycle,* Toronto: Office de la qualité et de la responsabilité en éducation.

Author (1999). *Pour un enseignement efficace des mathématiques et des sciences: Mathématiques de niveau avancé et sciences, 12^e année/CPO,* Toronto: Office de la qualité et de la responsabilité en éducation.