The Relevance of TIMSS for Policy Making in Education

David F. Robitaille
University of British Columbia

The K–12 education sector is big business in Canada, as it is in virtually every country in the world. The overall expenditure on education in this country is in excess of 40 billion dollars annually, making education the second largest area of public expenditure, accounting for somewhat more than 20 percent of all public expenditures. Canada spends a larger proportion of its gross national product on education than any other industrial country, with the exception of some of the Scandinavian countries.

For the last decade and more, we have heard repeated demands for accountability in all areas of public expenditure, including education. Taxpayers and their elected representatives, as well as educators, researchers and policymakers want to be assured that the investments they have made and are continuing to make in the K–12 education system are sound. They want evidence to show that the publicly funded education system is providing our children with the kinds of knowledge, skills and attitudes that will enable them to fulfill their potential and take up their place as fully contributing members of society.

Over the years, a great deal of time, money and effort have been devoted to a search for factors that make a difference in education. In spite of this effort, very little is known in a systematic way about "what works" in education. We do not have good quantitative information about what makes schools successful, or whether tax dollars for education are being well spent. Nor do we have good information about how educational variables are linked to health and well-being, or about the nature of the relationships between students' success in school and their subsequent success in the labor market.

Indeed for a long time—perhaps beginning with the Coleman Study (Coleman, 1966) in the United States—it has been widely believed that what schools do has little impact on student outcomes, especially when compared to the impact of home- and society-based variables. More recent analyses (Suter, 2000) indicate, on the other hand, that between-country differences in how students are taught and in what they are taught ("opportunity to learn") may well account for much more of the variance in student outcomes than has previously been believed.

One of the problems we face in this connection is the lack of agreement across stakeholder groups about how to quantify excellence in education: which variables should be included in an analysis of school success or failure. Some "report cards on schools" seem to select variables more on the basis of how easily available the data are rather than on any sort of reasoned analysis of the importance of those variables or their centality as part of the mission of our schools.

The thesis of this paper is that studies such as the Third International Mathematics and Science Study (TIMSS)[i] have made and will continue to make a unique an important contribution to the development of our understanding of "what works" in education. The focus of TIMSS is on classrooms, teachers and curriculum as the three foci of educational research which seeks to increase our understanding of what schools can do to help students succeed. Because of this it is very important that Canadian schools, teachers, and students continue to participate in these studies and that substantial resources be devoted to in-depth analyses of the results of those studies.

Comparative International Studies of Education

Canadians are not alone in their concern about students' achievement and the implications of that performance for the future. Governments around the world have shown an interest in assessing what is learned in school, particularly in mathematics and science. This reflects a growing consensus that scientific literacy and economic productivity are tightly linked. It is now widely believed that the failure of a nation to educate its work force threatens that nation's ability to keep pace economically in the international marketplace. It has, therefore, become very important for nations to know more about the performance of their school systems.

People all over the world share the same sorts of concerns, and they too are determined to take steps to ensure that the children in their schools are given the tools that they need to lead fruitful lives in this increasingly technological age. For example, in a paper on the impact of IEA studies in their country, the authors (Monseur and Brusselmans-Dehairs, 1997, p. 13) said the following:

> On a recurrent basis, the IEA surveys have revealed wide variations in pupil performance in Belgian schools in both mathematics and science. This is an important issue for policymakers, particularly when assessing evaluation policies and practices. Although Belgium compared rather well internationally on mathematics in the last comparative study of mathematics and science education, a careful analysis of the TIMSS data is necessary. Situated in the heart of Europe, Belgium cannot ignore international comparisons that might help avoid excessive discrepancies in regard to its partners.

Valid international comparisons among systems of education are extremely difficult to make because of the numbers of variables involved, because of our lack of sophisticated ways of accounting for the influence of many of those variables, and because of the fact that some of the variables are beyond the influence and control of the educational system. Torsten Husén, the founder of the International Association for the Evaluation of Educational Achievement (IEA), once described international comparative studies of education as exercises in comparing the incomparable. In saying this, his intention was not to indicate that international comparisons and international studies were pointless activities. On the contrary, he was saying that international studies were important precisely because the comparisons were so difficult to make, and yet so important. Education systems, as products of particular cultures, differ from one another in a number of fundamental ways, and one needs to guard against the temptation to make oversimplified generalizations or comparisons.

The results of well designed international comparative studies of education can contribute significantly to the debates that swirl constantly around questions of educational excellence and around calls for increased accountability. Findings from an international study such as TIMSS can, and should be, used to inform those debates.

Comparisons among educational systems are not new; but, until fairly recently, they were mainly reports written by individuals who had travelled and visited schools in other countries. In the second half of the twentieth century, undoubtedly aided by rapid developments in computer technology and data analysis, interest in more quantitative approaches to educational comparisons grew.

The major "players" in the international comparisons arena in education have been the Educational Testing Service (ETS), the OECD AND IEA. ETS conducted two rounds of its International Assessment of Educational Progress in the 1980s, but has not been involved in any other studies of this kind since then. The OECD runs the ongoing PISA (Programme for International Student Assessment) studies. IEA has conducted studies in many curricular areas, but in this paper we are focusing on those that focus on mathematics and science.

The IEA studies, particularly TIMSS, differ from those conducted by either ETS or OECD in some fundamental respects, and these differences should be taken into consideration in the process of making decisions about which studies to participate in. One area of difference is whether national samples selected for participation should consist of intact classrooms of students in schools or samples of students selected from classrooms. A second area of difference has to do with whether the samples should be selected on the basis of age or grade level.

Arguments for and against these two questions are discussed in the planning process for every international study; and, as is frequently the case in such matters, there really is no correct answer. The designers of each study have to make design decisions on the basis of what they hope to learn from the study.

The goal of TIMSS and of other IEA studies is to assist countries in the task of improving the teaching and learning of mathematics and science in all of the participating countries. The goal is not simply to rank countries on the basis of students' achievement, but rather to use the data collected to contribute to the long, slow process of improvement. In order to do this, it must be possible to link curriculum with instructional practices and both of these with student outcomes. The best way to do that is to use a design that involves the selecting of intact classes of students in a particular grade or range of grades.

The IEA approach is to use national samples of intact classrooms of students at a particular grade level (or, as in the case of TIMSS, pairs of adjacent grade levels). Both the ETS and OECD studies use samples of individual students selected on the basis of age. The SAIP program in Canada also uses this approach.

On the face of it, it would seem a lot easier, and perhaps more understandable to the average taxpayer, to focus on age rather than grade level. After all, everyone knows what a 13-year old is, but who knows what the equivalent of Grade 8 in Canada is in some other country. It is also true that children start school at different ages in some countries (although almost all of them now start at age 6, as we do). The trouble with age as a design parameter is that the results from an age-based sample tell us little or nothing that can be used in making recommendations about teaching or curriculum. In Canada, for example, 13-year olds can be found in three different grade levels and, in some cases, in as many as five grade levels. This makes it impossible to connect students' achievement to either what teachers do or what the curriculum is like. Selecting a sample of students rather than a sample of intact classrooms poses the same problem. Classrooms can be connected to particular teachers, to particular instructional processes and to a particular curriculum. A sample of, say, 13-year-old mathematics students within a given school would be connected to a fairly large number of different teachers, and it would therefore not be possible to establish linkages between students' achievement and any teacher-based variables.

IEA and TIMSS

IEA is a consortium of research institutes, ministries of education, universities and a number of individuals in more than 50 countries. The mandate of IEA is to conduct comparative studies of educational processes and their outcomes as a means of assisting policymakers, researchers, educators and other stakeholders in their attempts to improve the education of children in the K–12 education system internationally.

Over the almost 40 years of its existence, IEA has conducted more than 20 comparative studies, beginning with the First International Mathematics Study (Husén, 1967), and culminating most recently in a partial replication of the TIMSS in 1999. Each IEA has its own particular focus; and, over the years, IEA studies have focused on virtually every major area of the K–12 curriculum.

For the past twenty years or more, beginning with the Second International Mathematics Study (Robitaille and Garden, 1989), IEA studies have had curriculum as their central focus. This stems from an early recognition in IEA studies of the importance of opportunity to learn as a crucial element in any comparative investigation of students' achievement. By way of illustration, the diagram in FIGURE 1 summarizes the conceptual framework for TIMSS. Curriculum was to be examined from three viewpoints or perspectives: the curriculum as mandated at the system level (i.e. the intended curriculum); the curriculum as taught by teachers in classrooms (the implemented curriculum); and the curriculum as learned by students (the attained curriculum). The goal of the study was to examine each of these three components of the curriculum separately and, most importantly, to look for relationships among the three.

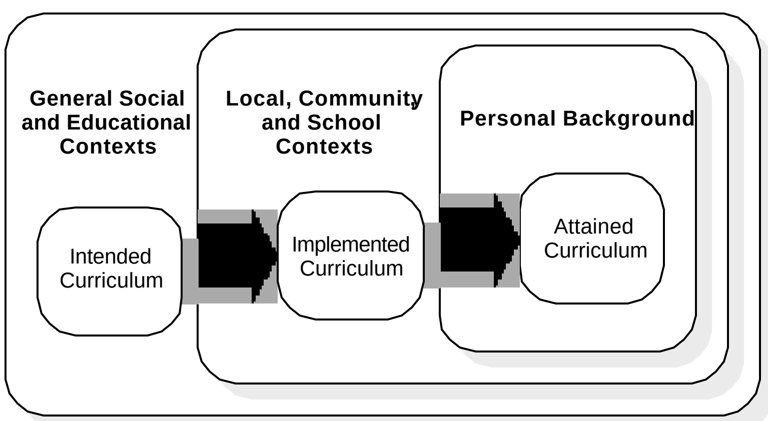

FIGURE 1. Conceptual framework for TIMSS.

Four research questions informed the design of TIMSS. They were the following:

1. How do countries vary in the intended learning goals for mathematics and science; and, what characteristics of educational systems, schools and students influence the development of those goals?
2. What opportunities are provided for students to learn mathematics and science; how do instructional practices in mathematics and science vary among nations; and what factors influence these variations?
3. What mathematics and science concepts, processes and attitudes have students learned; and what factors are linked to students' opportunity to learn?
4. How are the intended, the implemented and the attained curricula related with respect to the contexts of education, the arrangements for teaching and learning and the outcomes of the educational process?

Several international reports of the results from both the 1995 and 1999 rounds of TIMSS data collection have been published and they are available through the Website maintained by the International Study Center at Boston College (http://www.timss.org). Each participating country has published its own reports and the Canadian reports are available on the UBC-based Website at http://www.cust.educ.ubc.ca/wprojects/TIMSS/index.html.

The Canadian sample was structured in such a way as to make it possible to compare the Canadian national results with those from other countries, but also to compare the results from certain provinces with other countries. In both 1995 and 1999, the Canadian sample included schools and students from every province except Prince Edward Island, which declined to participate. On the other hand, five provinces selected samples that were large enough to make it possible to produce stable estimates of students' achievement at the provincial as well as the national level. In terms of the way the governance of education is structured in Canada, comparisons at the provincial level—whether between provinces within Canada or between provinces and other countries—are more likely to have an impact on teaching and learning in Canadian classrooms.

Selected Issues Emanating from the TIMSS Findings

The International Study Center at Boston College has not only published numerous reports of the findings from TIMSS, it has also released, in CD-ROM format, all of the data from the two rounds of data collection along with the appropriate documentation. All of the data are available through their Website at the URL noted earlier.

The good news is that there is considerable evidence available that TIMSS has had an impact on the teaching and learning of mathematics and science in many countries, and a volume containing country-by-country descriptions of that impact is available (Robitaille, Beaton & Plomp, 2000). The bad news, at least in Canada, is that very little seems to have been done to promote systematic secondary analysis of the TIMSS data to investigate relationships among curricula, instructional practices and student outcomes. Certainly the TIMSS–Canada data have been used by researchers in a number of areas, but not by many in educational research. The Ontario Ministry of Education did fund some follow-up analyses focusing on the Ontario data, but that appears to have been the only case of its kind so far.

This is unfortunate. The design of TIMSS makes it possible for policymakers, researchers and educators to investigate linkages between the intended, the implemented and the attained curricula. That kind of analysis is not possible based on other designs such as those employed in PISA and SAIP. We need to promote research based on the TIMSS data and we need to provide the resources needed to carry out that research.

One very important application of research based on TIMSS data is related to what Torsten Husén used to describe as "using the world as an educational laboratory". Many educational variables do not lend themselves to manipulation for research purposes. For example, no country would agree to have some children start school a year later than usual for a research study. Or again, no country would be likely to allow class size limits to vary for research purposes. But, those two variables—along with many others—are subject to some degree of variation across countries and this can be used to investigate the impact of those variables in a more naturalistic fashion. We can investigate what kinds of curriculum can be offered to students, what organizational patterns within classrooms seem to be most efficacious and a host of other questions that might be very difficult impossib le to do within any one country.

In the following paragraphs, a small number of policy-related issues that could be illuminated by reference to the TIMSS data are introduced. This is by no means an exhaustive list; it is intended to be illustrative of the kinds and range of discussions for which the TIMSS data would be useful. Others would undoubtedly identify other issues of equal or greater importance. This list consists mainly of issues directly related to policymaking in the realm of curriculum and instruction.

Organizing instruction for improved teaching and learning. The TIMSS databases include a great deal of information collected from teachers and students about how instruction in mathematics and science is organized and carried out. To date, little attention has been paid to investigating the relationships between various instructional practices and students' achievement or attitudes. There is every reason to believe that this data could make a significant contribution to our knowledge about teaching practices and their impact on students' learning. This is an area that needs attention.

Questionnaires to be completed by principals, teachers and students are integral parts of virtually every IEA study including TIMSS and the process of developing those instruments is an area where some improvement is still needed. Part of the reason that there has not been more research on this aspect of the TIMSS data may well be that the questionnaires do not contain enough of the kinds of items researchers would like to see. The way to solve that problem is by becoming involved in a study early enough to have a say in what goes into the instruments. This is true for the achievement tests as well as the questionnaires.

Effective schools. Some schools appear to be more successful than others. It may be that a higher proportion of their students complete graduation requirements. Or, perhaps, more of their students take advanced courses and are successful. Or again, perhaps students from those schools have more positive attitudes. Whichever criteria are used to operationalize the concept of what being an effective school means, what is it about those schools that make a difference?

The TIMSS data include a great deal of information about school climate, organizational patterns, students' and teachers' attitudes and opinions and other variables of interest in this domain. Whether one wanted to do this research solely on schools within Canada or focus on comparisons between Canadian schools and those in other countries of interest to Canadians, this is an area in which the TIMSS data could be extremely useful.

High-achieving students. The first TIMSS–Canada report (Robitaille, Taylor & Orpwood, 1996) identified the fact that, in many other countries, a much higher proportion of students were achieving at the highest levels than was the case in Canada. Similar findings have been identified in subsequent reports.

This situation should be a matter for urgent consideration by policymakers as well as other stakeholders in education. We should have as many high-achieving students proportionally as other countries do and they should be doing at least as well as their counterparts in those other countries.

A related issue has to do who participates in advanced courses in mathematics and physics at the Grade 12 (or equivalent) level and how many talented students drop out of either mathematics or science for one reason or another. Every industrialized country is currently dealing with a shortage of skilled personnel with good backgrounds in science and technology.

Disparities in achievement. Each of the TIMSS–Canada reports has compared the results obtained by Canadian students overall to those in other countries as well as to the results obtained by students in those provinces that oversampled. It is striking, if perhaps not surprising, to see how great the disparaties are among the provinces in this regard. Students in Québec and Alberta perform at almost the same level as those in the highest-achieving countries. Students in some other provinces do much less well.

This is a matter that needs serious attention both provincially and nationally. We need to get a better understanding of the reasons for these disparities and we need to do so soon. One thing that our examination of this situation to date has made clear is that these disparities are not strongly linked to differences in curriculum. We do not know yet why, say, the scores in Newfoundland and in Alberta are so different; but we know that it cannot be explained away by differences in curriculum between the two provinces.

Use of technology in mathematics and science. The TIMSS data show that Canadian schools, whether elementary or secondary, are well equipped with computers and related technology. Those data also show that, whatever else those computers and other items of equipment are being used for, they are not being used in mathematics and science classrooms either by teachers or by students. Unfortunately, that is about all the TIMSS data have to say in this regard.

While it may have been true until quite recently that teachers were not using computers to teach mathematics and science because of a dearth of good software, this is clearly not the case any longer. In both areas powerful new software is available, software that makes it possible to illuminate many difficult concepts in highly graphic and much more easily understandable ways. Apparently this message has not yet reached our schools.

Canadian sampling for TIMSS. The nature of the sample selected to represent Canada or any of the provinces in a study such as TIMSS, constrains the kinds of analyses that can be made once the data have been collected. In both rounds of TIMSS conducted so far, no attempt has been made to stratify the sample using criteria such as language of instruction, or publicly funded versus independent schools. This matter requires further discussion before any future study.

For TIMSS-95, countries were encouraged to select two classrooms per grade level in each participating school, in order that analysts interested in variance decomposition, for example, would be able to partition the variance into three components: between schools, between classrooms and between students. In the event, only two or three countries decided they could afford or wanted to have such a large sample. Canada was one of that latter group and this is another example of a decision whose implications need to be thoroughly discussed in connection with the next round of TIMSS which is scheduled for 2003.

Conclusion

TIMSS will focus on Grades 4 and 8 and there are indications that as many as 50 countries plan to participate. Of course, only a subset of those countries will be of interest for comparisons with Canada, but that group will include all of the other members of the G–8, almost all of the OECD countries and a large number of countries of interest to Canadians for historical and cultural reasons. As this paper is being written, there is still no assurance that there will be any Canadian participation in the project and this would be a great shame.

One of the things that TIMSS is concentrating on is an examination of trends in the teaching and learning of mathematics and science as well as trends in patterns of students' achievement. This is the kind of information that could be of great value to Canadian educators and policymakers and we can scarcely afford to ignore the opportunity to get that kind of information on a regular basis. We need to be able to see how we and others are moving and changing and we need to be able to interpret that information

Of course, funding and jurisdiction over education are issues. There is only so much money available for this kind of research and education is, after all, a provincial responsibility. However, TIMSS is a real bargain compared to the other studies mentioned in this paper that are going on in the same time frame. Canadian membership in IEA and our participation in TIMSS are coordinated nationally through the Council of Ministers of Education and all of the provinces (excluding PEI once again) have indicated their interest in participating in TIMSS–2003.

Other studies may help us understand how Canadian students' achievement and their attitudes compare to those of students in other countries. They may also be able to link students' achievement to a number of input variables. However, only TIMSS can help us improve our understanding of what kinds of curricula and what kinds of instructional practices are associated with the highest levels of student attainment in mathematics and science. And that's exactly what we need to know.

References

Husén, Torsten (Ed) (1967), *International Study of Achievement in Mathematics* (New York: John Wiley and Sons).

Monseur, C. and Brussellmans-Dehairs, C. (1997), Belgium. In D. F. Robitaille (ed.), *National Contexts for Mathematics and Science*

Education: An Encyclopedia of the Educational Systems Participating in TIMSS (Vancouver: Pacific Educational Press).

Robitaille, D. F. and Garden, R. A. (eds.) (1989), *The IEA Study of Mathematics II: Contexts and Outcomes of School Mathematics* (Oxford: Pergamon Press).

Robitaille, D. F., Taylor, A. R. and Orpwood, G. (1996), *TIMSS–Canada Report: Volume 1, Grade 8* (Vancouver: Department of Curriculum Studies, University of British Columbia).

---

[i] The first round of data collection for TIMSS took place in 1995, followed by a partial replication of the study in 1999. A third round is scheduled for 2003. The acronym, TIMSS, now stands for the Trends in Mathematics and Science Study.