2-2002

Predictors of National Differences in Mathematics and Science Achievement: Data From TIMSS for Eighth Grade Students

Erling E. Boe
University of Pennsylvania, boe@gse.upenn.edu

Henry May
University of Pennsylvania

Gema Barkanic
University of Pennsylvania

Robert Boruch
University of Pennsylvania, robertb@gse.upenn.edu

Recommended Citation

Boe, E. E., May, H., Barkanic, G., & Boruch, R. (2002). Predictors of National Differences in Mathematics and Science Achievement: Data From TIMSS for Eighth Grade Students. *CRESP Research Report,* Retrieved from http://repository.upenn.edu/gse_pubs/411

Predictors of National Differences in Mathematics and Science Achievement: Data From TIMSS for Eighth Grade Students

Abstract

There is widespread belief that national economic productivity is related to student performance in mathematics and science. With the advent in the 1960s of international surveys of student achievement in math and science, cross-national comparisons have become possible and nations have aspired to become "world class" in this respect. A major national policy issue in the U.S. and elsewhere is to identify and implement actions to attain and maintain a high level of student achievement in math and science in international comparisons.

The math and science project reported here was designed to capitalize on the potential for cross-national research with the Third International Mathematics and Science Study (TIMSS). TIMSS demonstrated that there are wide differences among nations in average student knowledge of math and science at several grade levels. Accordingly, a major research problem is to explain the sources of these national differences; that is, to identify the national-level variables that are the strongest predictors of national differences in average achievement scores. This problem was investigated to generate new research-based knowledge relevant to policy making about math and science education.

Disciplines

Education

PREDICTORS OF NATIONAL DIFFERENCES IN MATHEMATICS AND SCIENCE ACHIEVEMENT: DATA FROM TIMSS FOR EIGHTH GRADE STUDENTS[1]

Erling E. Boe, Ph.D.
Professor of Education

Henry May, B.S.
Research Assistant

Gema Barkanic, B.A.
Programmer Analyst

Robert F. Boruch, Ph.D.
University Trustee Professor of Education

Research Report No. 2002-TIMSS2

Center for Research and Evaluation in Social Policy
Graduate School of Education
University of Pennsylvania
Philadelphia, PA 19104

Revised: February 28, 2002

---

[1] Support for this research was provided by a grant (Award Number R215U980021) from the Fund for Improvement in Education (CFDA Number: 84-215U), Office of Reform Assistance and Dissemination, Office of Educational Research and Improvement, U.S. Department of Education; by a grant (Grant Number REC-9815112) from the Research on Education, Policy, and Practice Program (NSF 96-138), Division of Research, Evaluation, and Communication, Directorate for Education and Human Resources, National Science Foundation; and by the Center for Research and Evaluation in Social Policy, the Graduate School of Education of the University of Pennsylvania.

Executive Summary

There is widespread belief that national economic productivity is related to student performance in mathematics and science. With the advent in the 1960s of international surveys of student achievement in math and science, cross-national comparisons have become possible and nations have aspired to become "world class" in this respect. A major national policy issue in the U.S. and elsewhere is to identify and implement actions to attain and maintain a high level of student achievement in math and science in international comparisons.

The math and science project reported here was designed to capitalize on the potential for cross-national research with the Third International Mathematics and Science Study (TIMSS). TIMSS demonstrated that there are wide differences among nations in average student knowledge of math and science at several grade levels. Accordingly, a major research problem is to explain the sources of these national differences; that is, to identify the national-level variables that are the strongest predictors of national differences in average achievement scores. This problem was investigated to generate new research-based knowledge relevant to policy making about math and science education.

Research Methods

Using national-level data from TIMSS for 8th grade students, approximately 180 input variables were created of five types: student background and behavior; student attitudes, beliefs, and perceptions; instructional factors; school factors; and national demographic and economic factors. The relationships between these input variables and the national average math and science achievement scores from TIMSS were quantified by sophisticated multivariate statistical methods. Results were replicated with independent samples of 7th grade students.

Results and Policy Implications for Mathematics

Based on the best model from multivariate analysis, three national-level predictor variables emerged as having the largest effect on national average math achievement scores. These variables alone accounted for 84% of the total variability among nations in math achievement. The three predictor variables were:

1. National percentage of students who strongly agreed that they usually do well in math (referred to below as the "usually do well in math" variable); this variable was inversely related to national math achievement scores.
2. National percentage of students who strongly agreed that their mother thought it was important for them to be good at sports (referred to below as the "mother sports" variable); this variable was inversely related to national math achievement scores.
3. National percentage of students who strongly agreed that to do well in math they need good luck (referred to below as the "need luck in math" variable); this variable was inversely related to national math achievement scores.

Results from further analyses of each of these predictor variables suggested several implications for policies designed to improve national average math achievement.

Specifically, the findings with respect to the variable "usually do well in math" were complex and seemingly contradictory. This variable was found to be positively related to math achievement at the student level within nations, but was inversely related to achievement at the national level across nations. This could suggest to policy makers and others either of the following two opposite conclusions:

- Within-nation findings at the student level suggest the adoption of instructional policies that make it particularly easy and enjoyable for students to learn math, and that contribute to their particularly liking math and perceiving that they do especially well in math. Such policies will be seen as constructive because these student attitudes and perceptions are all associated with somewhat higher levels of math achievement when students are compared with each other. However, this conclusion based on within-nation findings would mislead policy makers in their efforts to improve a nation's average level of math achievement as shown by the results of cross-national analyses.

Alternatively,

- Cross-national findings with national-level variables suggest the adoption of instructional policies that do not make it particularly easy and enjoyable for students to learn math, and that do not contribute to their particularly liking math and perceiving that they do especially well in math. Such policies will be seen as constructive because these student attitudes and beliefs are in fact all associated with much lower levels of national average math achievement when nations are compared with each other.

Therefore, the research findings from the cross-national analyses suggest that sufficiently high standards of instruction and achievement should be set so that students do not experience learning math as particularly easy and enjoyable, and that students should be sufficiently challenged so that they do not particularly like math and do not perceive themselves as usually doing especially well in math. Accordingly, the important implication for educational policy from analyses of the variable "usually do well in math" is that:

- High national standards for teaching and learning math should be set in order to enhance international competitiveness in student achievement.

In contrast, the variable "mother sports" was found to be inversely related to math achievement in both within-nation and cross-nation analyses. The implications for social policy are that:

- The social status of educational achievement in math should be enhanced, visible and strong incentives for students to achieve at a high level should be created, and the status of, and rewards for, achievement in non-academic endeavors such as sports should be de-emphasized.

With respect to the variable "need luck in math" that was also found to be inversely related to math achievement both within nations and across nations, the implications for parents and for educational policy are that:

- Children should be inculcated in the concept that learning in school requires hard work instead of good luck, and a milieu should be created in which the work ethic is modeled and successful.

Results and Policy Implications for Science

Based on the best model from multivariate analysis, four national-level predictor variables emerged as having the largest effect on national average science achievement scores. These variables alone accounted for 85% of the total variability among nations in science achievement. The four predictor variables were:

1. School provides any special science enrichment activities (referred to below as the "science enrichment" variable); this variable was <u>positively</u> related to science achievement scores.

2. Capacity of school to provide instruction is affected "a lot" by a shortage or inadequacy of science laboratory equipment and materials (referred to below as the "laboratory equipment shortage" variable); this variable was <u>inversely</u> related to science achievement scores.

3. Hours students work on paid job per week (referred to below as the "student paid job" variable); this variable was <u>inversely</u> related to science achievement scores.

4. National percentage of students who strongly agreed that to do well in science they need good luck (referred to below as the "need luck in science" variable); this variable was <u>inversely</u> related to science achievement scores.

Results from further analyses of each of these predictor variables suggested several implications for policies designed to improve national average science achievement.

Specifically, the variable "science enrichment" was found to be positively related to science achievement both within nations and across nations. The implication for educational policy is that:

- Schools should be enabled, and required, to offer science enrichment activities.

With respect to the variable "laboratory equipment shortage" that was inversely related to science achievement both within nations and across nations, the implication for educational policy is that:

- Schools should be enabled, and required, to offer science instruction supported by adequate laboratory equipment and supplies.

With respect to the variable "student paid job" that was found to be inversely related to science achievement both within nations and across nations, the implications for social policy are that:

- Family economic status should be sufficiently enhanced so that 13-year old students do not need to work at a job to earn money. Alternatively, strengthen child labor laws prohibiting employment of 13-year old adolescents.

With respect to the variable "need luck in science" that was also found to be inversely related to science achievement both within nations and across nations, the implications for parents and for educational policy are that:

- Children should be inculcated in the concept that learning in school requires hard work instead of good luck, and a milieu should be created in which the work ethic is modeled and successful.

Conclusions

The primary results of this research, and their policy implications, were obtained from a cross-national analysis of relationships between input variables and achievement outcomes in math and science. In contrast, this research was not designed to explore the causes of differences in achievement among students within any nation, and the cross-national results should not be generalized to the student level.

Nonetheless, our findings are valid at the national level, and should be useful for policy making intent on improving a nation's relative standing in cross-national comparisons of math and science achievement. This research supports the following general conclusions:

- Educational factors represent one major source of differences among nations in academic achievement. Our interpretation of findings is that national average math achievement is related to the rigor of national standards for teaching and learning, whereas national average science achievement is related to the adequacy of science laboratories and to the provision of enrichment learning opportunities. Other than these educational factors, national differences in average math and science achievement scores do not seem to be related to national differences in the breadth and repetition of curricular content, particular instructional methods, or teacher qualifications. Such variables may be important determinants of achievement differences among students within nations, but do not account for average achievement differences across nations.

- Cultural factors represent another major source of differences among nations in average academic achievement. Cross-nationally, it appears that the emphasis mothers place on their children being good at sports is inversely related to student achievement in math; likewise, the amount of time students spend working at a paid job is inversely related to science scores.

- A particular student belief (i.e., needing good luck to do well in math and science) represents a third major source of national differences in average math and science scores. Cross-nationally, the strength of this attitude is inversely related to both math and science achievement.

- Because national average full-scale and sub-scale scores are highly correlated, similar findings should be expected if cross-national analyses were performed with TIMSS sub-scale scores in math and science (as contrasted with the comprehensive or full-scale scores used in our research).

TABLE OF CONTENTS

Page

Introduction

Policy Issue

There is widespread belief that national economic productivity is related to student performance in mathematics and science (e.g., see Peak, 1996, pg. 35, representing the view of the U.S. Department of Education). As stated by Rita R. Colwell, Director of the U.S. National Science Foundation, "It is critical that students in the United States achieve at high levels in mathematics and science. The position of the U.S. in the world economy, the continuing demand for well trained mathematicians and scientists, and the need for an informed citizenry able to make intelligent public-policy decisions about important economic, medical, and environmental issues all depend upon it" (Gonzales et al., 2000, p. vii).

With the advent in the 1960's of international surveys of student achievement in math and science, cross-national comparisons have become possible and nations have aspired to become "world class" in this respect.[1] A major national policy issue in the U.S. and elsewhere is to identify and implement actions to attain and maintain a high level of student achievement in math and science in international comparisons.

In addressing this issue, policy makers seek the best research available on the determinants of student achievement in math and science.[2] It is quite possible that some of the most useful research for this purpose will be obtained from cross-national analyses of educational systems, processes, and student background, and how these differences relate to variation in student achievement. TIMSS is a prime source of cross-national data to support such research.

---

[1] For example, U.S. public policy makers (state Governors, the U.S. President, and later the U.S. Congress) adopted a policy goal of U.S. students becoming "first in the world in mathematics and science achievement by the year 2000" (National Education Goals Panel, 1991).

[2] When the available research is inadequate, policy makers also seek to expand the research base, such as supporting the extensive research on math, science, and engineering education by the U.S. National Science Foundation.

Project Background

The math and science project reported here was designed to capitalize on the potential of research with TIMSS to inform educational policy making. Using TIMSS, research on math and science achievement can be conducted at two levels:

1. Student-level research conduced within nations, with students as the unit of analysis.
2. National-level research conducted across nations, with nations as the unit of analysis.

In pursuing national-level research with TIMSS, three main strategies can be followed:

1. Cross-national comparisons of educational inputs (e.g., characteristics of educational systems such as centralization of decision-making authority, level of educational funding per student, curriculum content, family background of students, etc.) based on univariate analyses,
2. Cross-national comparisons of educational outcomes (e.g., general achievement scores in math, subscale scores in algebra, scores in higher-order thinking, school completion rates, etc.) based on univariate analyses, and
3. Cross-national relationships between educational inputs and outcomes based on bivariate and multivariate analyses.

This project focused on the third strategy; viz., on identifying national-level input variables that are associated with national differences in average student achievement scores in math and science.

In view of the technical refinement that went into its design and the subsequent quality of data collection, TIMSS has been widely recognized as the largest and most sophisticated comparative education study accomplished to date. Among its many components are questionnaire surveys of various educational inputs (administered to schools, teachers, and students) and outcome measures of student achievement in mathematics and science.

Review of a wide variety of reports of analyses of TIMSS data (e.g., Beaton, Mullis, et al., 1996; Beaton, Martin, et al., 1996) reveals a major interest in the relative standing of participating nations in terms of average student achievement scores in mathematics and science as shown in Figures 1 and 2, respectively. Six nations (Japan, Germany, Hong Kong[3],

---

[3] The Peoples Republic of China did not participate in TIMSS. However, China was represented in TIMSS by Hong Kong, though, of course, the Hong Kong TIMSS sample is not representative of China in any respect.

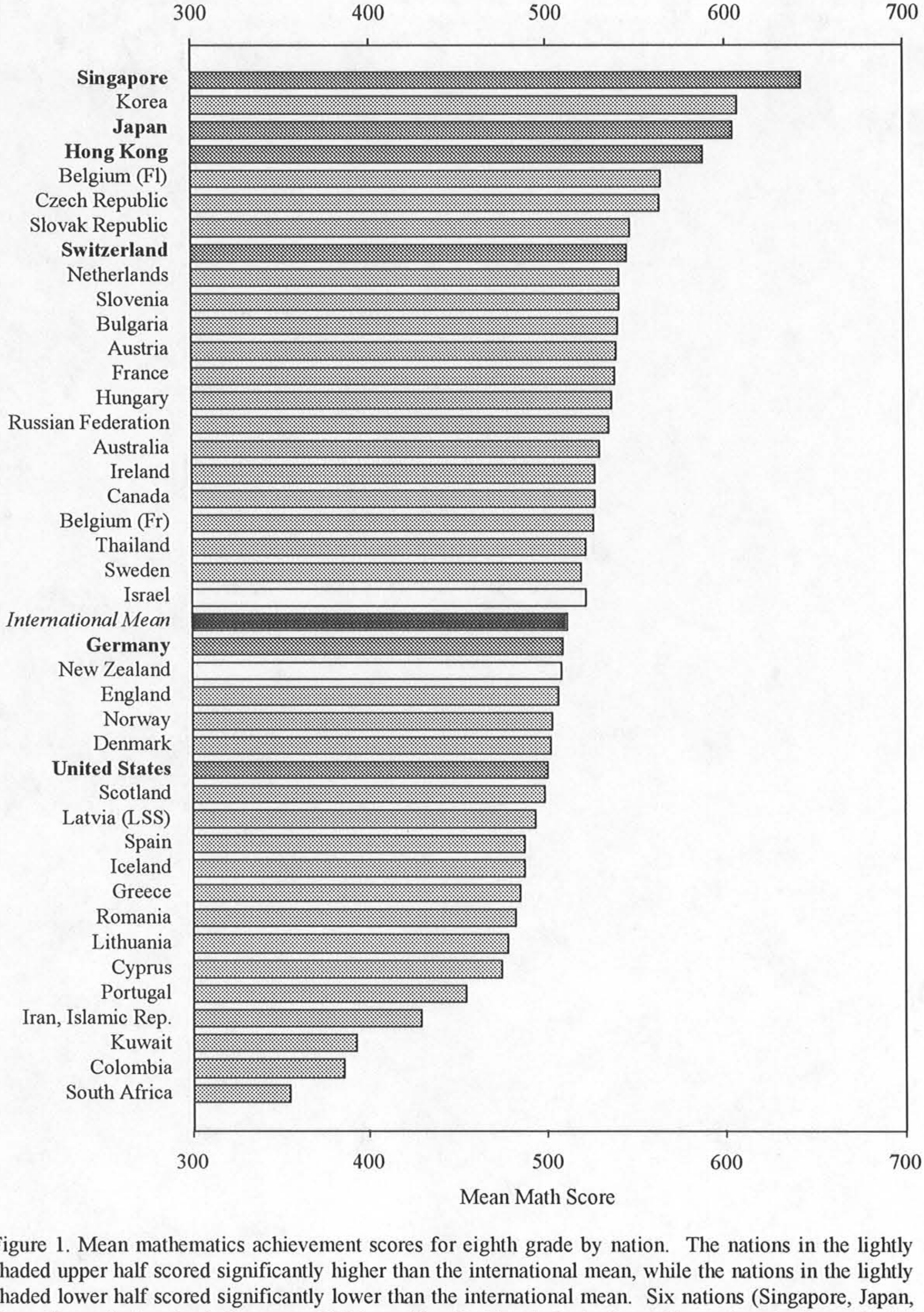


Figure 1. Mean mathematics achievement scores for eighth grade by nation. The nations in the lightly shaded upper half scored significantly higher than the international mean, while the nations in the lightly shaded lower half scored significantly lower than the international mean. Six nations (Singapore, Japan, Hong Kong, Switzerland, Germany, USA) are identified by dark shading. The nations in the unshaded center section are not statistically different than the international mean. Data source: TIMSS.

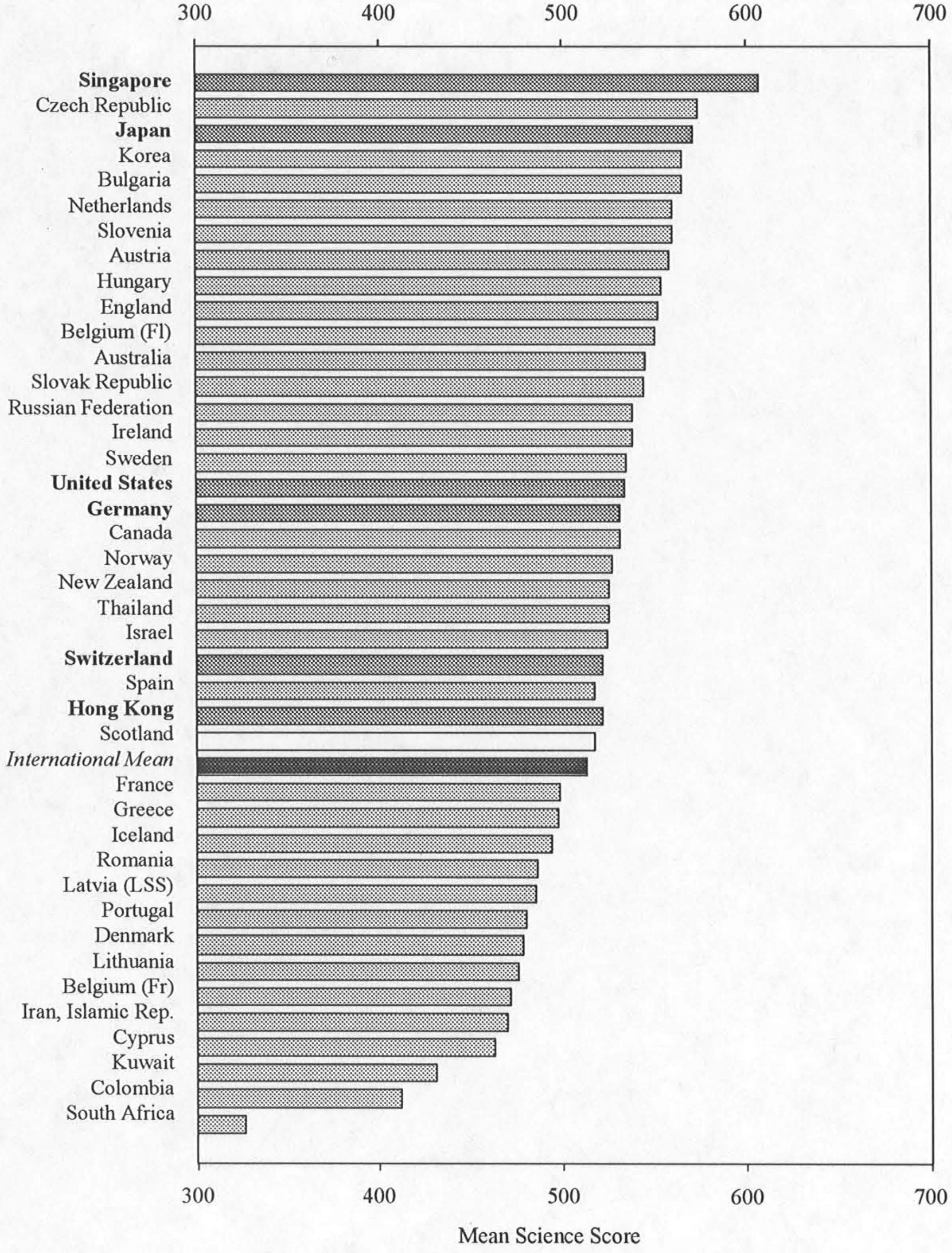


Figure 2. Mean science achievement scores for eighth grade by nation. The nations in the lightly shaded upper half scored significantly higher than the international mean, while the nations in the lightly shaded lower half scored significantly lower than the international mean. Six nations (Singapore, Japan, Hong Kong, Switzerland, Germany, USA) are identified by dark shading. The nations in the unshaded center section are not statistically different than the international mean. Data source: TIMSS.

Singapore, Switzerland, and the United States) are highlighted in these and subsequent figures to mark the position of three nations from the East and three from the West. In addition, these and other official TIMSS reports have reported extensive cross-national data on educational input and student background variables (e.g., teacher certification requirements, class size, homework assigned, educational level of students' parents, etc.).

In spite of the great interest in understanding the causes of national differences in student achievement in math and in science, none of the official reports of TIMSS issued to date have provided quantification of the cross-national relationships between educational input and achievement outcome variables. This observation applies to the numerous reports by the TIMSS International Study Center (e.g., Beaton, Mullis, et al., 1996; Martin, Mullis, Gregory, Hoyle, & Shen, 2000[4]), the U.S. National Research Center for TIMSS (e.g., Schmidt et al., 1999), and the National Center for Education Statistics (NCES), USDE (e.g., NCES, 2000).

In contrast with these official TIMSS reports, the results of a microeconomic study of TIMSS by a secondary analyst (Woessmann, 2001) have been reported. Woessmann included in his analysis several national-level variables, and found that nationally-centralized examinations and national levels of private school enrollment were positively related to national differences in mean math scores, after controlling for multiple input variables. Though Woessmann examined a few national-level relationships, the emphasis of his study was at the student level. Thus, it appears that the capacity of the largest and most sophisticated comparative education survey (i.e., TIMSS) to provide quantitative evidence about sources of national differences in achievement has yet to be exploited.

Research Problem

As seen in Figures 1 and 2, the results of TIMSS achievement testing showed that there existed large national differences in average student knowledge of math and of science. Accordingly, a major research problem is to explain the sources of these national average differences; that is, to identify the national-level variables that are the strongest predictors of variation among nations in average achievement scores. This approach should be distinguished from within-nation studies that attempt to explain variation in math and science achievement at

---

[4] This report included results of multivariate analyses of predictors of achievement scores within several nations with schools as the unit of analysis. However, it did not report cross-national analyses of achievement with nation as the unit of analysis.

the student level. This research was designed to exploit the ability of TIMSS to yield new information about national differences in achievement by performing a cross-national analysis, and thereby expand the research base directly relevant to the policy issue of how to "attain and maintain a high level of student achievement in math and science in international comparisons."

Research Methods

An overview of the analysis methods used in this research are presented in this section; a more detailed presentation is found in Appendix B: Technical Methods. The analyses focused on the national level, separately for math achievement and for science achievement. Except for subsidiary multilevel analyses, the primary unit of analysis was the "nation" instead of student, classroom, or school. The TIMSS sample nations is described in Appendix B.

Data Sources

The main data source for this research was nationally-representative TIMSS data for eighth and seventh grade students from nations that participated fully at each of these two grade levels in 1995 (41 nations for grade 8; 39 nations for grade 7). Most of the educational and student data were collected by extensive questionnaires administered to students, teachers, and schools (the latter completed by principals of schools). In addition, TIMSS (Beaton, Mullis, et al., 1996, Table 4) and other data sources were used to provide national-level economic, social, demographic, and educational variables.

National-Level Outcome Variables

The outcome (or dependent) variables were student achievement measures; specifically, national mean scores in math and in science at grade 8 (as shown in Figures 1 and 2, respectively) and at grade 7. The comprehensive (i.e., full scale) scores for math and science (as distinguished from subscale scores for algebra, geometry, physics, biology, etc.) were used in these analyses.

National-Level Predictor Variables

The national-level predictor (i.e, input or independent) variables included a large number of potential predictors of the national mean scores in math and in science. The selection of predictor variables for analysis was based on several considerations. A considerable number were first selected because of their relevance to educational policy, practice, and/or theory (e.g., class size and length of the school year). Next, many variables were selected because they represented major dimensions (e.g., curriculum characteristics and teacher qualifications) of the comprehensive TIMSS conceptual framework (Martin & Kelly, 1996, Figure 5.4). Others were selected for analysis because they were included in reports issued by the TIMSS International Study Center (e.g., teacher age and student perceptions of importance mothers attach to doing well in sports, as reported by Beaton, Mullis, et al., 1996). A number of variables were added because of special interest (e.g., student time watching TV and aspects of school safety), while others were selected because of the presumed importance of certain educational practices (e.g., the frequency with which teachers meet to plan curriculum and instructional methods, and student time spend on extra lessons out-of-school). In addition, major national variables (e.g., public expenditures on education per capita) were obtained from other sources (as listed in Boe et. al., 2001, Appendix A) to augment the TIMSS data base.

Most of these national-level predictor variables were created by aggregating individual responses to the TIMSS student, teacher, and school questionnaires to the national level. As an example of aggregation to produce a national-level variable, the mean of student reports of the number of hours worked per day for pay was computed for all students responding at a particular grade level in a particular nation.

Five main types of national-level predictor variables were analyzed:

1. **Student Background and Behavior**: Examples of the 28 predictor variables of this type are student age, possession of a computer at home, hours per week worked at a job for pay, time spent on homework.

2. **Student Attitudes, Beliefs, and Perceptions**: Examples of the 37 predictor variables of this type are student agreement with: need to memorize math to do well, liking math, mother thinks it important to do well in sports.

3. **Instructional Factors**: Examples of the 54 predictor variables of this type are years of teacher experience, teacher professional development, breadth of curriculum prescribed.

4. **School Factors**: Examples of the 46 predictor variables of this type are the number of instructional days in the school year, average class size, shortage of instructional equipment.

5. **National Demographic and Economic Factors**: Examples of the 13 predictor variables of this type are national levels of life expectancy, gross domestic product per capita.

The complete list of the 178 predictor variables included in the analysis is presented in Appendix A; questionnaire item numbers and wording, and sources of each is found in Boe, et al. (2001, Appendix A). To obtain the overall national average for each predictor variable in the first four categories, questionnaire responses items within each nation were aggregated.

Analysis Design

All analyses were performed cross-nationally (separately for math achievement and for science achievement), to examine the relationships between (a) each national-level predictor variable and (b) each national-level outcome variable (the math and science achievement scores). Accordingly, the main unit of analysis at grade 8 was the 41 nations in TIMSS (rather than students, classrooms, or schools).

Bivariate Analyses. Given that TIMSS data were obtained through a large-scale survey (as distinguished from field experimentation or other methods), relationships among predictor-outcome variables were most appropriately analyzed with correlational techniques that provided for the quantification of such relationships at the national level. Since these were exploratory analyses of 178 particular predictor-outcome relationships separately for math and science, replication of statistically-significant bivariate correlations found with eighth grade students was examined in the independent national samples of seventh grade students. Upon computing product-moment correlations and after exclusion of selected predictor variables for multicolinearity, 35 national-level predictor variables for mean math scores were identified. Similarly, 38 national-level predictor variables for mean science scores were identified.

Multivariate Analyses. The predictor variables identified through the bivariate analyses were then used in multivariate regression, using both ordinary least squares (OLS) (Neter, Kutner, Nachtschiem, & Wasserman, 1996). Using this method, it was possible to identify the main national-level predictor variables that emerge as the best predictors of national differences in average math and science scores, while other potential predictors are shown simultaneously to be less powerful. Since such analyses of TIMSS have not been performed before and were therefore exploratory, these analyses with eight grade students were replicated with independent national samples of seventh grade students from TIMSS. A comparison of the grade 8 results with grade 7 results was made to determine whether the results were similar in independent samples of students (i.e., were replicated across samples).

Multilevel Analyses. The multivariate prediction models found with OLS method were next recomputed by the hierarchical linear modeling (HLM) method (Raudenbush & Bryk, 2002) for two reasons. First, to determine whether the national-level models identified by the OLS method would be confirmed by a different analytic method. Second, to examine the predictive power of the best national-level models of achievement scores at the classroom and student levels.

Results and Discussion of Mathematics Achievement

Math Score Variability at the National Level

As seen in Figure 1, there were large differences in national mean math scores at grade 8. The frequency distributions of math achievement scores for individual students, along with their means and standard deviations, within each of three nations from the East and three from the West, are illustrated in Figure 3. In comparison with the international average of 41 nations, the mean math score for the U.S. was significantly below average, for Germany not significantly below average, and for significantly above average for Switzerland, Hong Kong, Japan, and Singapore (Beaton, Mullis, et al., 1996).

The analysis and interpretation of national differences in mean math scores is only worthwhile if a considerable percentage of total variability in individual math scores in the TIMSS student sample (i.e., across all students from all TIMSS nations combined) is attributable to the national level. If, for example, the percentage of national-level variability were small, there would be little value in performing comparative research to explain why nations differ in math achievement.

The results of HLM analyses of full-scale math score variance (MSV) at the eighth grade level indicated that 30% of total math score variability was attributable to differences among nations. This supports the need for comparative education research to explain the role of national-level predictor variables in producing variation in national mean math scores (the dependent variable in the analyses reported here). The remaining MSV is attributable to the student level (50%) and the classroom and school levels combined (20%). Unfortunately, the TIMSS sampling design precludes precise subdivision of the 20% of MSV allocated to the classroom and school levels.

Bivariate Predictor-Outcome Variable Relationships for Math

Although these differences among nations in mean math scores are of interest in their own right, the main purpose of this research was to explain why nations differ in math achievement; that is, to identify the national-level variables that are the strongest predictors of variation across nations in mean math scores. By contrast, others have reported analyses of mean math scores

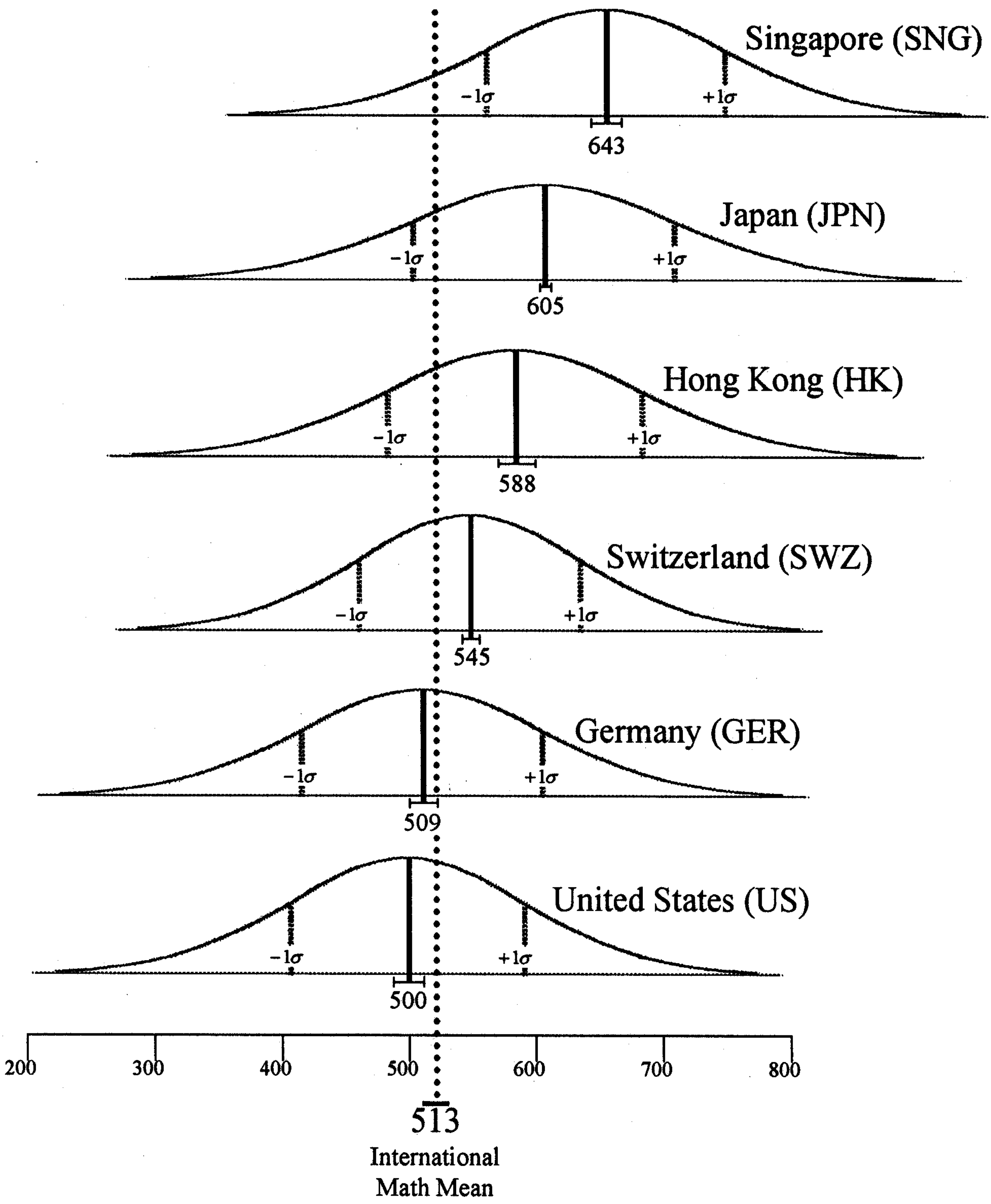


Figure 3. Eighth grade mathematics achievement from TIMSS: Frequency distributions of achievement scores for each of six nations with national mean scores and associated 95% confidence intervals (⊢⊣), as compared with the international mean score.

at the school level within nations (Martin, et al., 2000) and at the student level within nations (Baker, Goesling, & LeTendre, in press).

The results of cross-national correlations between 178 predictor variables and mean math achievement are shown in Appendix A separately for independent samples of seventh and eighth grade students. As seen there and as listed in Table 1, these bivariate analyses yielded 35 statistically significant correlations for eighth grade students that were also statistically significant (i.e., replicated) in independent samples of seventh grade students. (A few variables that were significant predictors at both grade levels were excluded due to coliniarity with selected variables.)

From Table 1, it can be seen that there was quite a marked difference in the large number of nationally aggregated student attitude/belief variables were strongly associated with national mean math scores when contrasted with the few, and relatively weak, relationships found with instructional variables. This difference occurred even though a large number and variety of instructional variables were examined.

Also of importance are the many predictor variables seen in Appendix A that were not correlated cross-nationally with mean math scores. Many of these are of special interest because the TIMSS Conceptual Framework (Martin & Kelly, 1996, Figure 5.4, p. 5-8) identifies a variety of policy-relevant student, teacher, curriculum, instruction, and educational system variables that are presumed to be linked to variation in mean student achievement scores across nations. Our analyzes test whether these presumed relationships are observed cross-nationally in TIMSS data. Listed below are a number of general categories of predictor variables shown in the TIMSS Conceptual Framework to be related to variation in math scores cross nationally, followed by a relevant example of a particular nationally-aggregated variable provided by TIMSS that was not related cross-nationally to differences in mean math scores:

1. Student background: Percentage of foreign born students
2. Student activities: Percentage of students who took extra lessons in math outside school
3. Student attitudes: Percentage of students who strongly agree that to do well in math you need lots of hard work studying at home
4. Teacher qualifications: Years of teaching experience
5. Teacher status: Percentage of full-time teachers

Table 1. *Thirty-five National-Level Variables Used to Predict Eighth-Grade Mean Math Achievement in Multivariate Analyses Across 41 Nations (missing nations not imputed)*

	Bivariate Correlation with National Mean Math		
Predictor Variable	r	p	N
Student Background and Behavior			
Average number of books in the home	.35	.03	40
Percent of students who have a desk, dictionary, calculator, and computer at home	.41	.01	40
Average hours per day students do jobs at home	-.58	.00	41
Average hours per day students watch television and videos	.31	.05	41
Percent of students who skipped class at least once in past month	-.52	.00	37
Average hours per week students work at a paid job	-.51	.00	40
Student task persistence (i.e., the national percent of items asked in student background questionnaires actually answered)	.79	.00	41
Student Attitudes and Beliefs			
Percent of students who strongly agree that: I usually do well in math at school	-.60	.00	41
Percent of students who strongly agree that: it is important to do well in math	-.39	.01	41
Percent of students who strongly agree that: I need to do well in math to get the job I want	-.69	.00	41
Percent of students who strongly agree that: it is important to be placed with high achieving students	-.58	.00	40
Percent of students who strongly agree that: math is an easy subject	-.76	.00	40
Percent of students who strongly agree that: I enjoy learning math	-.76	.00	40
Percent of students who strongly agree that: you need talent to do well in math	-.49	.00	40
Percent of students who strongly agree that: you need good luck to do well in math	-.72	.00	40
Percent of students who strongly agree that: you need to memorize the textbook or notes to do well in math	-.43	.01	40
Percent of students who strongly agree that: students often neglect their school work	-.61	.00	40
Percent of students who strongly agree that: mother thinks it is important to be good at sports	-.76	.00	39
Instructional Variables			
Percent of students who work in small groups at least once in a while	-.39	.01	41
Percent of students whose math teachers report that uninterested students limit how they teach "quite a lot" or more	-.34	.05	39
Average hours per week teachers meet with students outside the formal school day	.32	.05	38
Average hours per week teachers meet with parents outside the formal school day	-.36	.05	38

Note: Data from the Third international Mathematics and Science Study (TIMSS)

Table 1 (Continued). *Thirty-five National-Level Variables Used to Predict Eighth-Grade Mean Math Achievement in Multivariate Analyses Across 41 Nations (missing nations not imputed)*

Predictor Variable	Bivariate Correlation with National Mean Math		
	r	p	N
School Variables			
Average number of instructional days per year	.38	.02	37
Percent of students who finish the school year in the same school	.50	.00	39
Percent of students whose principals report that their school's capacity to provide instruction is affected "a lot" by a shortage or inadequacy of science laboratory equipment and materials	-.42	.01	38
Percent of students whose math teachers report that inadequate physical facilities limit how they teach	-.44	.01	38
Number of boys repeating grade	-.60	.00	38
Percent of school age cohort in secondary schools	.36	.02	39
Demographic and Economic			
Log. Population density (per square km)	.44	.00	41
Average years of life expectancy	.40	.01	41
GDP (1990) per capita adjusted for PPPI	.49	.00	41
Percent of GDP in manufacturing	.30		34
Percent of the GDP in agriculture, hunting, forestry, and fishing	-.41	.01	38
Percent of labor force employed in industry	.35	.03	38
Percent of labor force employed in agriculture	-.38	.02	38

Note: Data from the Third International Mathematics and Science Study (TIMSS)

6. Teacher social organization: Number of times teachers meet with other teachers to plan curriculum and approaches to teaching
7. Teacher pedagogic beliefs: Importance for students to memorize formulas and procedures
8. Intended curriculum: (a) Breadth or (b) Repetition
9. Instructional activities: Percentage of teaching time based on textbook
10. System characteristics: Student age at start of compulsory schooling

These 10 variables merely illustrate the many predictor variables of policy relevance that are not associated with math achievement scores cross-nationally. Though too numerous to list and discuss here, all predictor variables analyzed are found in Appendix A.

The Multivariate Model for Mathematics Achievement

Though the bivariate relationships of each of the predictor variables with national mean math scores as shown in Appendix A are of considerable interest individually, of greater interest and importance are those included in the best multivariate prediction models. Therefore, multivariate analysis were used to quantify the simultaneous association of predictor variables with math achievement and to identify the predictor variables with the largest effects. The best multivariate model[5] (as generated by the OLS method) is shown in the top panel of Figure 4. As can be seen, three variables (all based on questionnaire responses of students) accounted for 84% of the total variability among nations in mean math achievement. These variables were:

1. National percentage of students who strongly agreed that they usually do well in math (referred to below as the "usually do well in math" variable).
2. National percentage of students who strongly agreed that their mother thought it was important for them to be good at sports (referred to below as the "mother sports" variable).
3. National percentage of students who strongly agreed that to do well in math they need good luck (referred to below as the "need luck in math" variable).

---

[5] The "best multivariate model" is defined by three criteria: an R^2 higher than the R^2 of any competing model; statistically significant parameter estimates for each variable in the model; and parsimony (i.e., a model comprised of the fewest predictor variables, such that the addition of one more predictor variable would not substantially increase R^2).

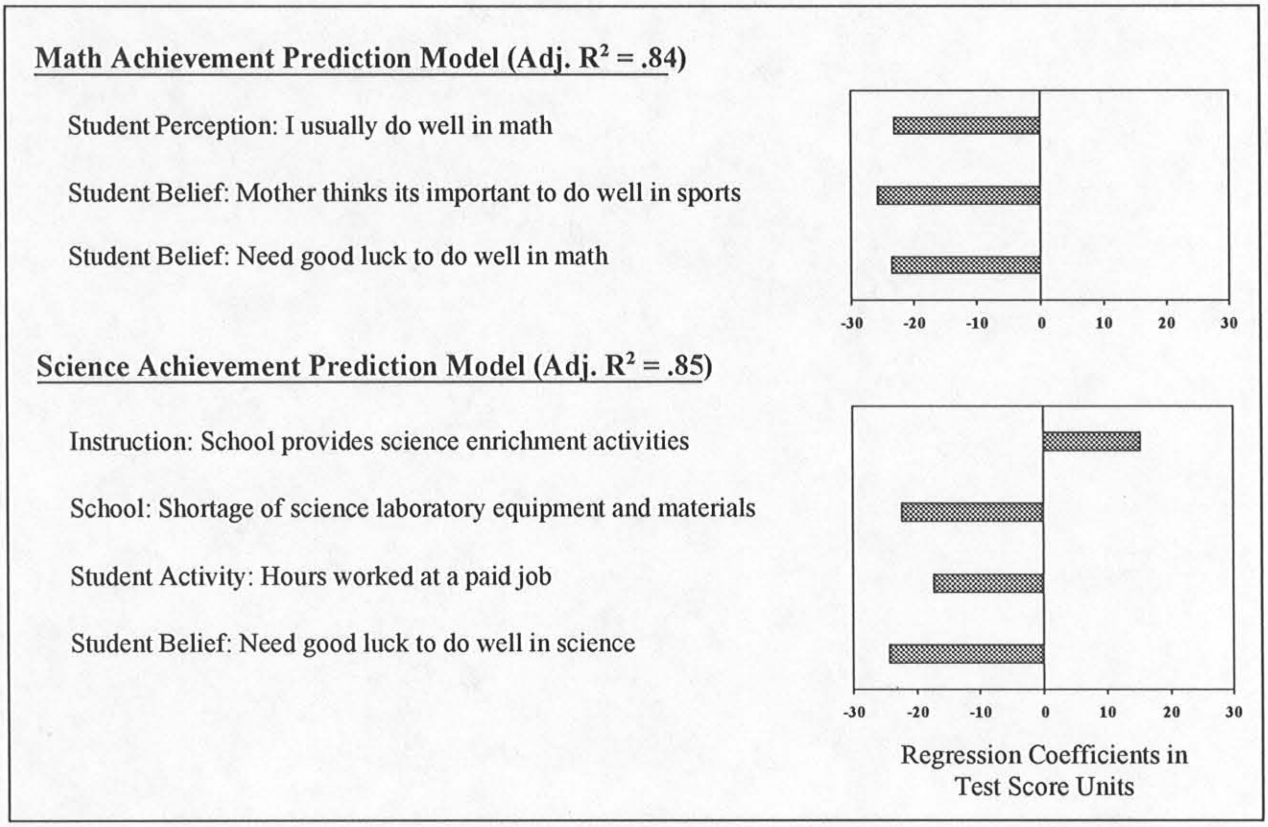


Figure 4. Results from multivariate analyses: The best prediction. Multivariate models of national mean mathematics and science achievement scores at the eighth grade. The regression coefficients for each variable are reported in achievement test score units, and represent the change in achievement for each standard deviation of a predictor variable. The adjusted R^2 for the mathematics model is .84, while the adjusted R^2 for the science model is .85. Data Source: TIMSS.

The power of this three-variable model to predict national mean math scores with little error is illustrated in Figure 5 as indicated by the close relationship between the observed national mean math scores and the national mean math scores predicted by the model. Note also that the three Western nations are arrayed in order along the regression line, whereas the high-scoring three Eastern nations are bunched at the high end of the predicted national scores.

When this three-variable prediction model was recomputed with independent national samples at the 7th grade, it accounted for 83 percent of the national-level variability in mean math achievement scores. Thus, the best cross-national multivariate model at grade 8 for predicting math achievement was replicated with independent national samples of students from grade 7.

In interpreting the results of these analyses at the national level, it is important to recognize that they are likely to be quite different from results of analyses that could be conducted at the student level. Therefore, generalizations between levels should not be made without clear evidence of their validity. To do so without such evidence is to commit the "ecological fallacy."

Multilevel Analysis for Math

In the HLM math model, the sample reduction after listwise deletion of cases with missing data was modest. Approximately 13% of the total sample of 147,505 students, 11% of the total sample of 5,879 schools, and 10% of the total sample of 41 nations were dropped for this analysis.

The HLM and OLS methods of analysis produced estimates of variance explained at the national level (approximately 84%) by the three-variable math model shown in Figure 4. The regression coefficients for each national-level predictor in the HLM model were also similar to those from the OLS method. Although not estimated using OLS, the estimate of variance explained within nations by the group-centered student-level versions of the predictor variables in the HLM model was much smaller (only 8.9%) than at the national level. Thus, the powerful three-variable national-level model was replicated using HLM, although the predictive power of this three-variable model within nations was much less.

Supplementary Analyses of Main Math Predictor Variables

The results of the cross-nation multivariate analysis of math achievement may seem surprising and even counter-intuitive, considering that none of the educational variables (e.g.,

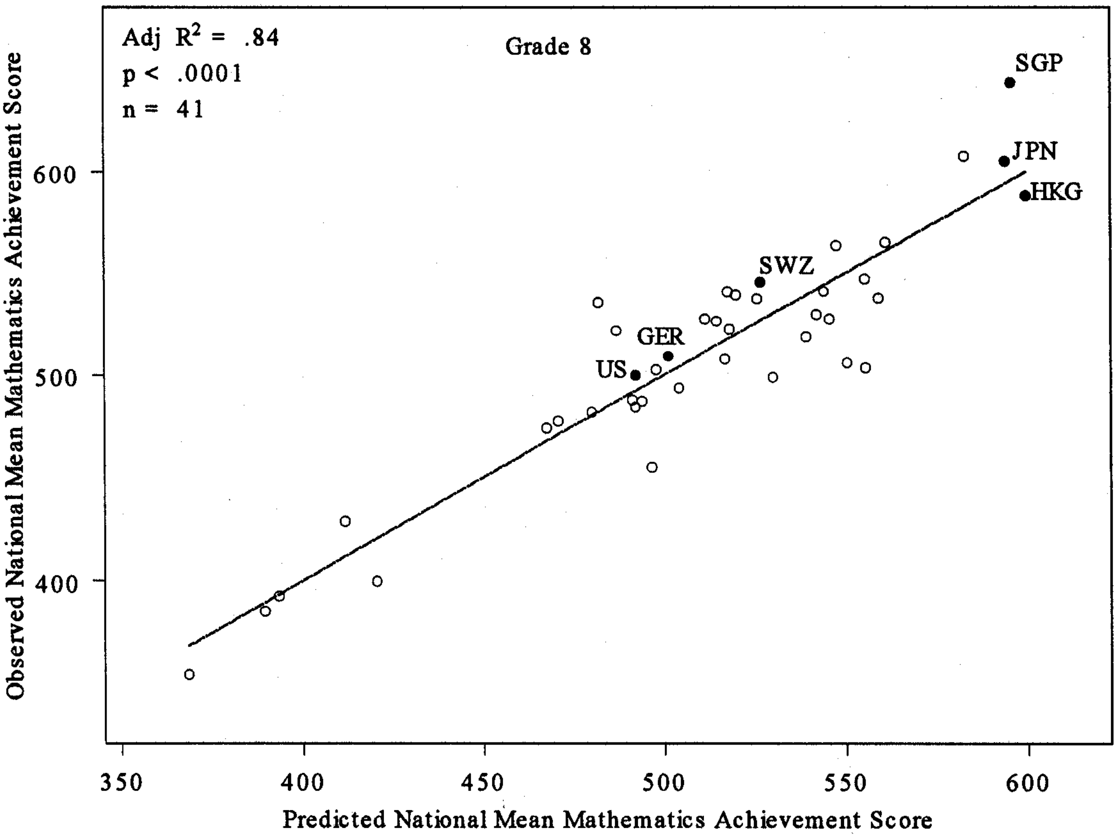


Figure 5. Observed national mean mathematics achievement scores as a function of predicted mean mathematics achievement scores for 41 nations. Predictions are based on the best multivariate model using three predictor variables. Six observations are highlighted: Germany (GER), Hong Kong (HKG), Japan (JPN), Singapore (SGP), Switzerland (SWZ), and the United States of America (US). Model adjusted $R^2 = .84$. Data Source: TIMSS

as average instructional days per year) or national wealth variables (e.g., GDP per capita) emerged as major predictors. Accordingly, we performed additional analyses to explore further each of the three predictor variables that emerged as main effects in the best model of national mean achievement in math as shown in Figure 4. A policy implication is drawn from each variable as a predictor of national-level math achievement.

Additional Analyses of the Predictor Variable "Usually Do Well in Math"

The observed inverse relationship between the variable "usually do well in math" and mean math achievement at the national level (as seen in Figures 4 and 6) indicates that the stronger the perception of doing well by students of a nation, the lower the nation's mean achievement scores. To investigate this puzzling cross-national finding, we performed within-nation analyses at the student level separately for each of the 41 nations. As might be expected, we found that within each nation the strength of student perceptions of doing well in math was in fact directly related to math achievement scores. Therefore, the contrast between the inverse relationship observed cross-nationally (top panel of the scatterplot in Figure 6) and the direct relationship observed within nations (bottom panel of the scatterplot in Figure 6, for all nations combined) is quite dramatic. The most plausible and reasonable interpretation of the within-nation results is that, within their own national "frame of reference," student perceptions of the level of their performance are positively related to their actual performance (i.e., students have a general idea of how well they do in math).

To understand the inverse relationship seen cross-nationally for the variable "usually do well in math" with math achievement, we correlated this variable cross-nationally with all other predictor variables, to identify which other predicator variables correlated most highly with it. The variables that were found to be related most strongly cross-nationally to "usually do well in math" were:

1. National percentage of students who strongly agreed that math is an easy subject (referred to below as the "math is easy" variable) (correlation with "usually do well in math" = .67)
2. National percentage of students who strongly agreed that they enjoyed learning math (referred to below as the "enjoy learning math" variable) (correlation with "usually do well in math" = .65)
3. National percentage of students who strongly agreed that they liked math a lot (referred to below as the "like math" variable) (correlation with "usually do well in math" = .60)

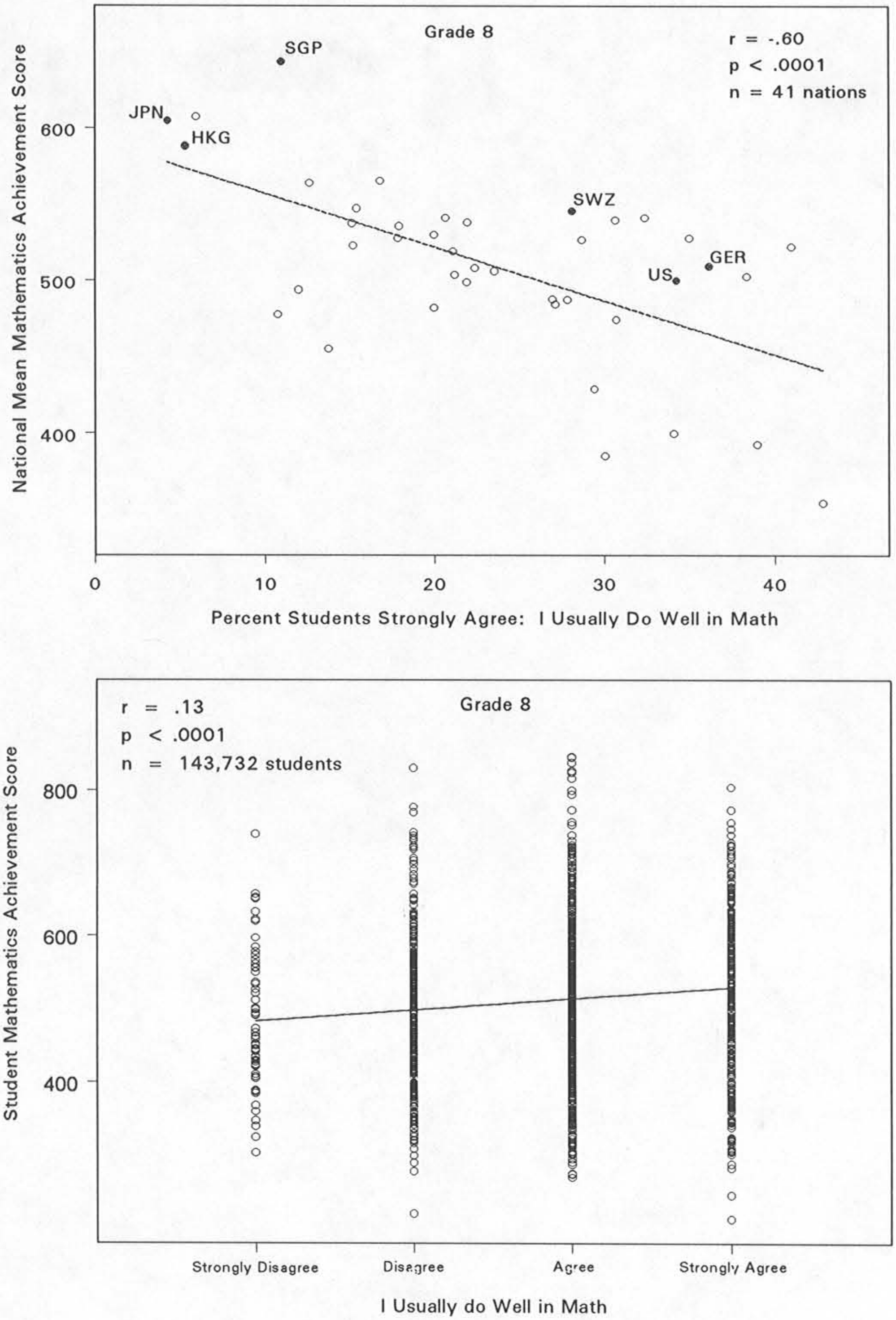


Figure 6. Comparison of cross-national (top panel) versus within-nation (bottom panel, pooling students from 41 nations) bivariate relationships between math achievement and student agreement with the statement "I usually do well in math." Top panel is based on imputed national data. Each data point in the bottom panel represents 100 students. Six observations are highlighted in the top panel: Germany (GER), Hong Kong (HKG), Japan (JPN), Singapore (SGP), Switzerland (SWZ), and the United States of America (US). Data Source: TIMSS.

As with the variable "usually do well in math" (as shown in Figure 6), further within-nation analyses of the three variables most strongly associated with it indicated that each one was inversely related to math achievement cross-nationally but was directly related to achievement at the student level within each nation.

These complex and seemingly contradictory findings may suggest to policy makers and others either of the following two opposite conclusions:

- Within-nation findings at the student level suggest the adoption of instructional policies that make it particularly easy and enjoyable for students to learn math, and that contribute to their particularly liking math and perceiving that they do especially well in math. Such policies will be seen as constructive because these student attitudes and perceptions are all associated with somewhat higher levels of math achievement when students are compared with each other.

Alternatively,

- Cross-national findings with national-level variables suggest the adoption of instructional policies that do not make it particularly easy and enjoyable for students to learn math, and that do not contribute to their particularly liking math and perceiving that they do especially well in math. Such policies will be seen as constructive because these student attitudes and beliefs are in fact all associated with much lower levels of national average math achievement when nations are compared with each other.

Thus, if a policy goal is to improve the national average math achievement, then the cross-national results are paramount because they are driven by national level variables instead of by student-level variables within nations. Accordingly:

- Research findings from within-nation analyses can mislead policy makers by suggesting that instruction making math particularly easy and enjoyable will enhance the average national level of student achievement in math,

Whereas,

- Research findings from cross-national analyses suggest that sufficiently high standards of instruction and achievement should be set so that students do not experience learning math as particularly easy and enjoyable, and that students should be sufficiently challenged so that they do not particularly like math and do not perceive themselves as usually doing especially well in math.

We therefore hypothesize that variation in national standards for teaching and learning math underlies observed national differences both in achievement and in student perceptions and attitudes; i.e., the higher the national standards, the higher students score on the math achievement tests, and the less students perceive that they do especially well in math, the less they find math easy, the less they enjoy learning math, and the less they like math. To the extent this hypothesis is correct, the student perceptions and attitudes examined here are indicators of a fundamental quality of national education systems (i.e., standards). The important implication for educational policy of the variable "usually do well in math" is that:

- High national standards for teaching and learning math should be set in order to enhance international competitiveness in student achievement.

To be effective, of course, this will require an educational system that can deliver high quality instruction consistently.

Additional Analyses of the Predictor Variable "Mother Sports"

Analyses similar to the above were repeated with the variable "mother sports." This variable was also found to be associated inversely with math achievement cross-nationally (see Figures 4 and 7). Similarly, inverse relationships were observed also within each of the 39 nations available for this analysis. This consistency of inverse association in the cross-national analysis (top panel of Figure 7) and the within-nation analysis (bottom panel of Figure 7, for all nations combined) might well be expected.

Further analyses were performed to identify which other predicator variables correlated most highly cross-nationally with the variable "mother sports." The two variables that correlated most highly with "mother sports" were:

1. National percentage of the GDP in agriculture, hunting, forestry, and fishing (correlation with "mother sports" = .51), and
2. National percentage of teachers who say inadequate physical facilities limit teaching a great deal (correlation with "mother sports" = .46).

Thus, it appears that the variable "mother sports" is more prominent in developing nations with agriculture-based economies and less investment in education facilities.

In view of these consistent inverse relationships observed both within and across nations, we hypothesize that parental emphasis on non-academic activities of their children, such as sports, diverts attention and effort away from learning math. The implications of this hypothesis for social policy are that:

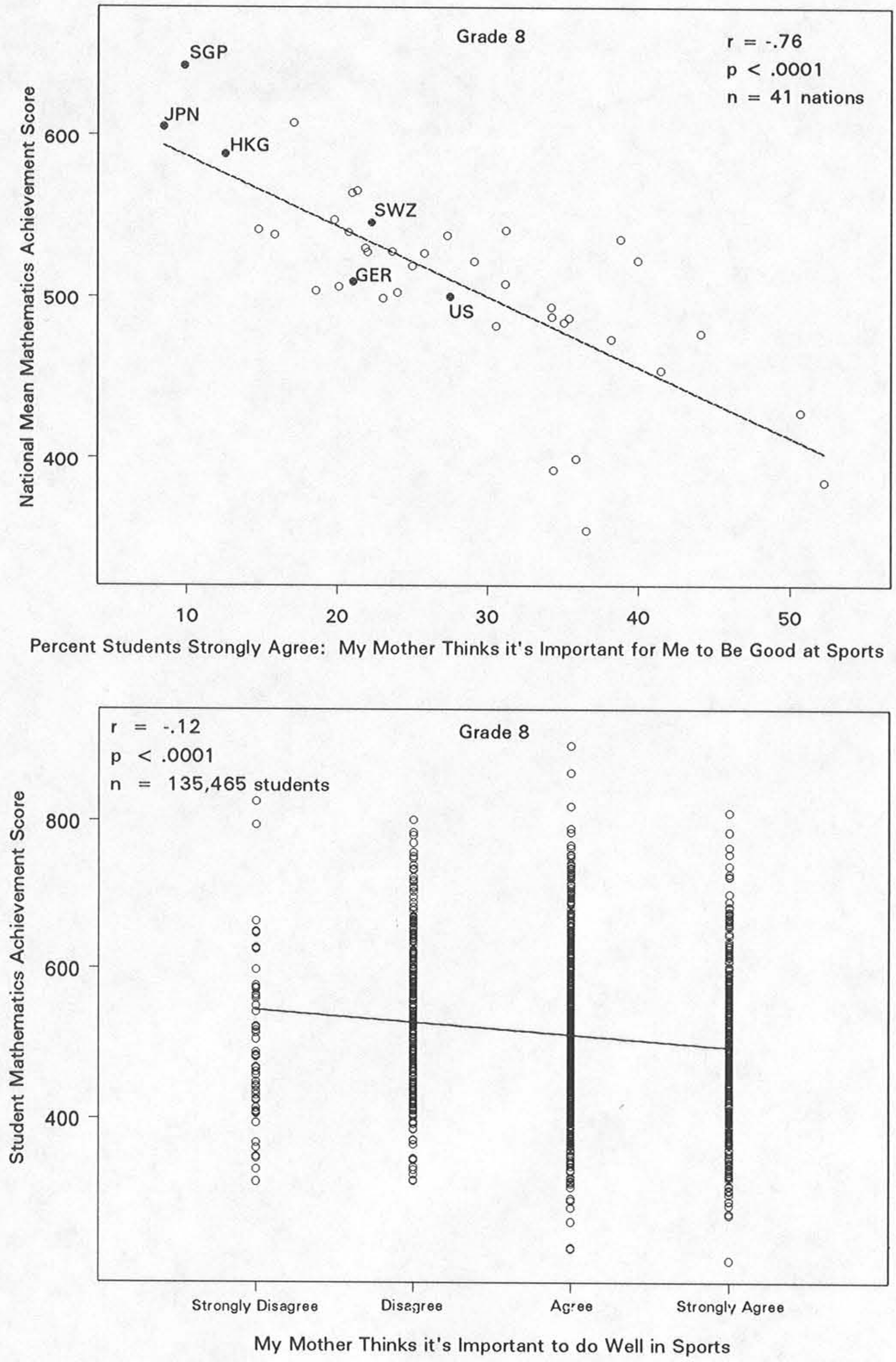


Figure 7. Comparison of cross-national (top panel) versus within-nation (bottom panel, pooling students from 39 nations) bivariate relationships between math achievement and student agreement with the statement “My mother thinks it’s important to do well in sports.” The top panel is based on imputed national data. Each data point in the bottom panel represents 100 students. Six nations/areas are highlighted in the top panel: Germany (GER), Hong Kong (HKG), Japan (JPN), Singapore (SGP), Switzerland (SWZ), and the United States of America (US). Data Source: TIMSS.

- The social status of educational achievement in math should be enhanced, visible and strong incentives for students to achieve at a high level should be created, and the status of, and rewards for, achievement in non-academic endeavors such as sports should be de-emphasized.

Additional Analyses of the Predictor Variable "Need Luck in Math"

The variable "need luck in math" was also found to be inversely associated with math achievement cross-nationally (see Figures 4 and 8). Similarly, inverse relationships were observed also within each nation. This consistency of inverse association in the cross-national analysis (top panel of Figure 8) and the within-nation analysis (bottom panel of Figure 8, for all nations combined) might well be expected.

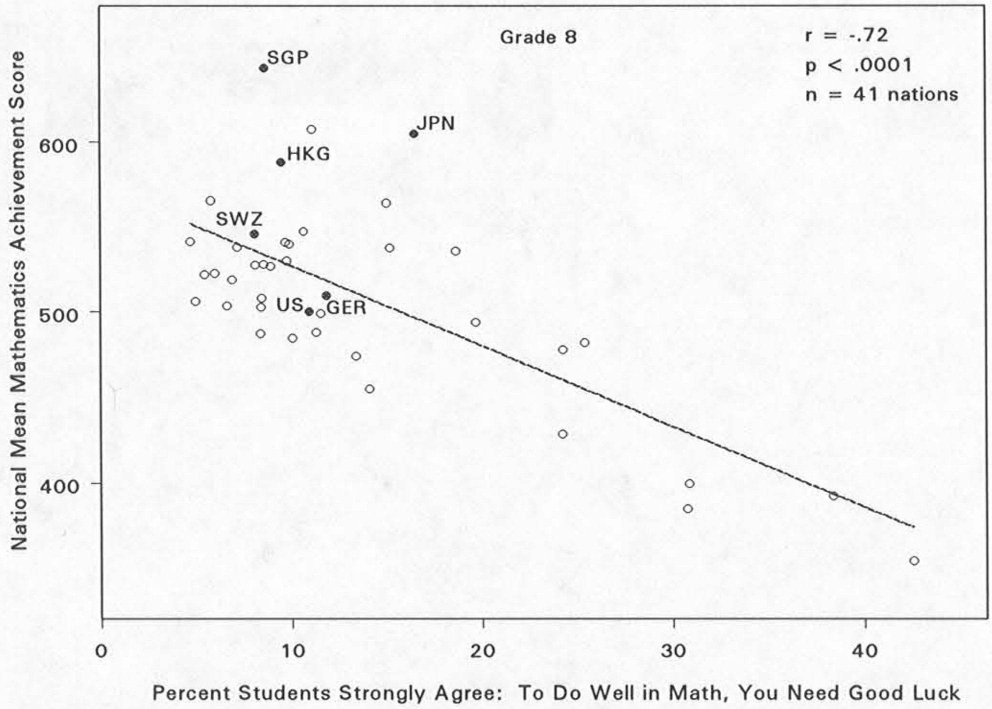
As with the "usually do well math" and "mother sports" variables, we identified other predicator variables that correlated most highly cross-nationally with the "need luck in math" variable. The two variables that correlated most highly with "need luck in math" were:

1. The national percentage of students who strongly agreed that math is an easy subject (correlation with "need luck in math" = .50)
2. The national percentage of students who strongly agreed that they would like a job that involved using math (correlations with "need luck in math" = .49)

Thus, it appears that the "need luck in math" variable (an unrealistic belief) is associated most strongly with two other variables that seem to represent unrealistic beliefs regarding math being easy and liking a job involving math. These belief's seem unrealistic because nations with students who tend to score high on all these variables tend to score low on math achiveemnt.

In view of the consistent inverse relationships of "need luck in math" with math achievement scores that are observed both within and across nations, we hypothesize that student beliefs that they need good luck to do well in math detracts from their making the effort required to actually do well. The implications of this hypothesis for parents and educational policy are that:

- Children should be inculcated in the concept that learning in school requires hard work instead of good luck, and a milieu should be created in which the work ethic is modeled and successful.



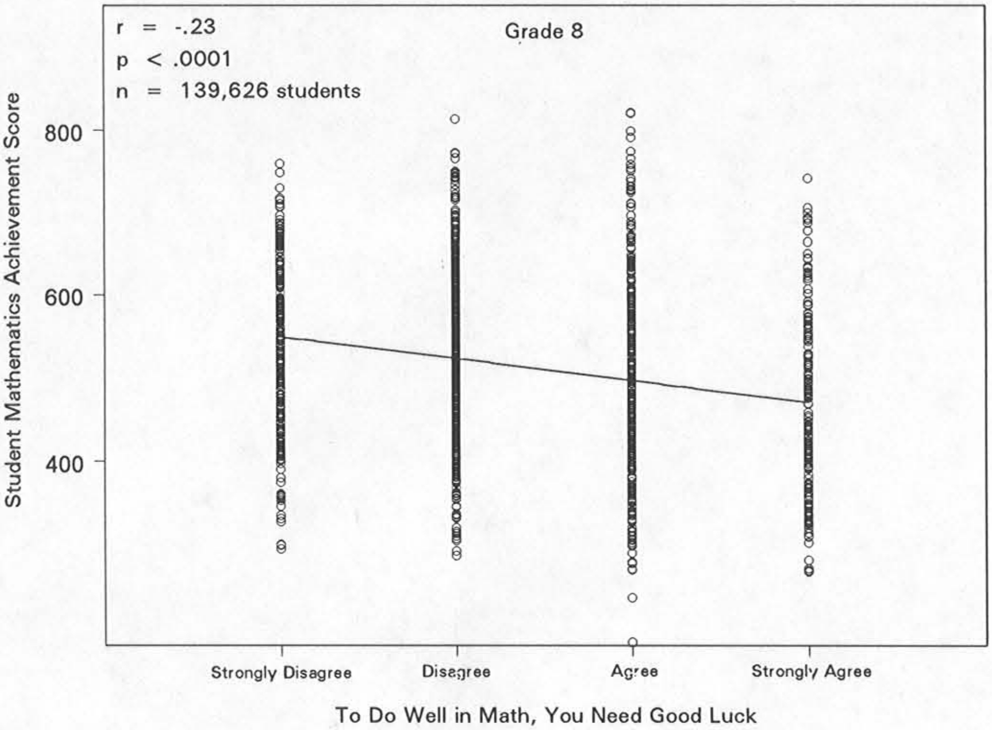


Figure 8. Comparison of cross-national (top panel) versus within-nation (bottom panel, pooling students from 40 nations) bivariate relationships between math achievement and student agreement with the statement "To do well in math, you need good luck." The top panel is based on imputed national data. Each data point in the bottom panel represents 100 students. Six observations are highlighted in the top panel: Germany (GER), Hong Kong (HKG), Japan (JPN), Singapore (SGP), Switzerland (SWZ), and the United States of America (US). Data Source: TIMSS.

Math Prediction Model Profile for Six Nations

Based on the results of the best multivariate model for math, a three-variable profile for predicting national average math scores is presented in Figure 9 for each of three nations from the East and three from the West. As seen in Figures 6, 7, and 8, all three predictor variables were inversely related to math achievement cross-nationally. Therefore, Singapore gained points on all three variables because its students did not believe that they usually did well in math, did not believe their mothers thought it important to do well in sports, and did not believe luck was needed to do well in math.

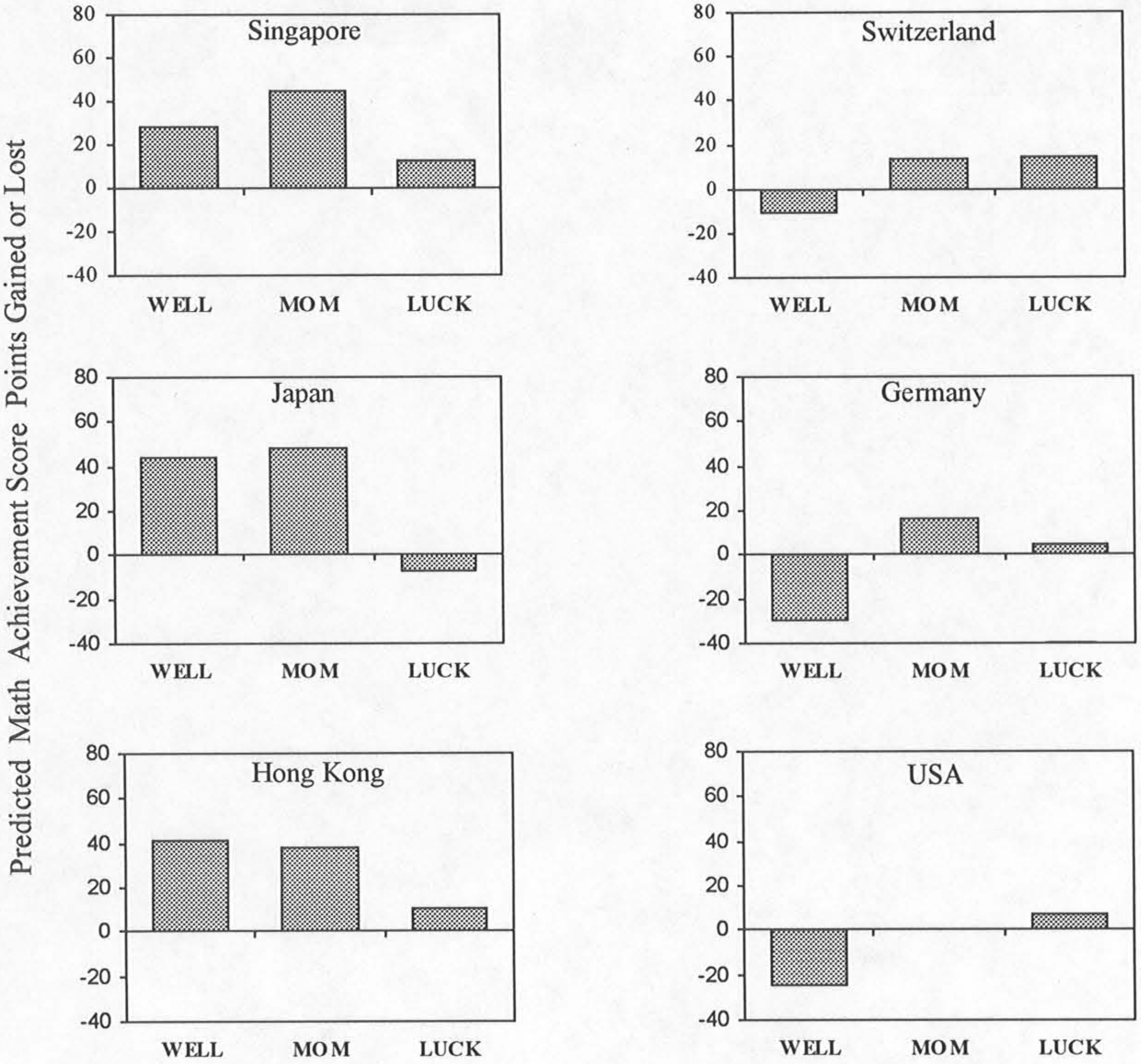
Perhaps the most interesting comparisons that can be made in Figure 9 are between the three nations from the East and the three nations from the West in terms of the degree to which students from each nation perceived that they usually do well in math. As seen in Figure 6, math performance decreased as the perception of doing well increased (an inverse relationship). In the three Eastern nations (Singapore, Japan, and Hong Kong), this belief is well below the international average as reflected in the prediction points gained by the "Well" variable. In contrast, this belief by students from all three Western nations (Switzerland, Germany, USA) was above the international average as reflected in the prediction points lost because of the "Well" variable. As stated in the previous section, we interpret variation in national average student perceptions of doing well in math as indicative of differences in national standards for teaching and leaning math (such that the stronger the national average student perception of doing well, the lower the national standards).

Results and Discussion of Science Achievement

Science Score Variability at the National Level

As seen in Figure 2, there were large differences in national mean science scores. As reported by Beaton, Martin et al. (1996), the mean science scores for five of the six reference nations (Germany, Japan, Singapore, Switzerland, and the U.S.) highlighted in this report were significantly above the international average, while the mean score for Hong Kong was not significantly higher. Singapore was higher, at a statistically significant level, than that of all nations participating in TIMSS.

The analysis and interpretation of national differences in mean science scores is only worthwhile if a considerable percentage of total variability in individual science scores in the



WELL: Students believe they usually do well in math (low belief gains points).

MOM: Students believe mother thinks it is important to do well in sports (low belief gains points).

LUCK: Students believe they need good luck to do well in math (low belief gains points).

Figure 9. Predicted math achievement points gained (+) or lost (-) as a function of each of three predictor variables included in the best multivariate model for math. Points gained, or lost, are measured from the international mean math achievement score of 510. Data Source: TIMSS.

TIMSS student sample (i.e., across all students from all TIMSS nations combined) is attributable to the national level. If, for example, the percentage of national-level variability were small, there would be little value in performing comparative research to explain why nations differ in science achievement.

The results of HLM analyses of full-scale science score variance (SSV) at the eighth grade level indicated that 20% of total science score variability was attributable to differences among nations. This supports the need for comparative education research to explain the role of national-level predictor variables in producing variation in national mean science scores (the dependent variable in the analyses reported here). The remaining SSV is attributable to the student level (64%) and the classroom and school levels combined (17%). Unfortunately, the TIMSS sampling design precludes precise subdivision of the 80% of total SSV allocated to the student, classroom, and school levels.

Bivariate Predictor-Outcome Variable Relationships in Science

Although these differences among nations in mean science scores are of interest in their own right, the main purpose of this research was to explain why nations differ in science achievement; that is, to identify the national-level variables that are the strongest predictors of variation across nations in mean science scores. By contrast, others have reported analyses of mean math scores at the school level within nations (Martin, Mullis, et al., 2000).

The results of cross-national correlations between 178 predictor variables and mean science achievement are shown in Appendix A separately for independent samples of seventh and eighth grade students. As seen there and as listed in Table 2, these bivariate analyses yielded 38 statistically significant correlations for eighth grade students that were also statistically significant (i.e., replicated) in independent samples of seventh grade students. (A few variables that were significant predictors at both grade levels were excluded due to coliniarity with selected variables.)

From Table 2, it can be seen that there was quite a marked difference in the large number of nationally aggregated student attitude/belief variables were strongly associated with national mean science scores when contrasted with the few, and relatively weak, relationships found with instructional variables. This difference occurred even though a large number and variety of instructional variables were examined.

Table 2. *Thirty-five National-Level Variables Used to Predict Eighth-Grade Mean Science Achievement in Multivariate Analyses Across 41 Nations (missing nations not imputed)*

	Bivariate Correlation with National Mean Science		
Predictor Variable	r	p	N
Student Background and Behavior			
Average number of books in the home	.36	.02	39
Percent of students who have a desk, dictionary, calculator, and computer at home	.43	.01	39
Average hours per day students spend doing math and science homework out-of-school	-.46	.00	39
Average hours per week students spend taking extra lessons in science out-of-school	-.37	.02	40
Average hours per day students do jobs at home	-.44	.00	40
Average hours per day students read a book for enjoyment	-.32	.05	40
Average hours per day students watch television and videos	.39	.01	40
Percent of students who skipped class at least once in past month	-.48	.00	36
Average hours per week students work at a paid job	-.36	.02	39
Student task persistence (i.e., the national percent of items asked in student background questionnaires actually answered)	.68	.00	40
Student Attitudes and Beliefs			
Percent of students who strongly agree that: I usually do well in science at school	-.45	.00	40
Percent of students who strongly agree that: it is important to do well in science	-.47	.00	40
Percent of students who strongly agree that: it is important to do well in the language used in school	-.42	.01	40
Percent of students who strongly agree that: I need to do well in science to get into the school I prefer	-.40	.01	40
Percent of students who strongly agree that: it is important to be placed with high achieving students	-.60	.00	39
Percent of students who strongly agree that: science is an easy subject	-.68	.00	40
Percent of students who strongly agree that: they enjoy learning science	-.74	.00	40
Percent of students who report that: they like science a lot	-.48	.00	40
Percent of students who strongly agree that: you need talent to do well in science	-.53	.00	39
Percent of students who strongly agree that: you need good luck to do well in science	-.67	.00	39
Percent of students who strongly agree that: mother thinks it is important to be good at sports	-.67	.00	38
Instructional Variables			
Percent of students whose science teachers report that a high student/teacher ratio limits how they teach "quite a lot" or "a great deal"	-.32	.05	39
National requirement of teaching practice or experience for teacher certification	.08	.63	40
Average hours per week teachers spend on administrative tasks	.46	.00	38

Note: Data from the Third international Mathematics and Science Study (TIMSS)

Table 2 (Continued). *Thirty-five National-Level Variables Used to Predict Eighth-Grade Mean Science Achievement in Multivariate Analyses Across 41 Nations (missing nations not imputed)*

Predictor Variable	Bivariate Correlation with National Mean Science r	p	N
<u>School Variables</u>			
Percent of students whose principals report that their schools' offer any special enrichment activities in science	.31	.06	39
Average number of instructional days per year	.37	.02	37
Percent of students who finish the school year in the same school	.51	.00	39
Nation administers curriculum-based external exit examinations in science	.41	.01	39
Percent of students whose principals report that their school's capacity to provide instruction is affected "a lot" by a shortage or inadequacy of science laboratory equipment and materials	-.53	.00	38
Percent of students whose science teachers report that inadequate physical facilities limit how they teach	-.42	.01	38
Number of boys repeating grade	-.61	.00	38
<u>Demographic and Economic</u>			
Average years of life expectancy	.31	.05	40
Public expenditures on education per capita	.28	.09	37
Number of scientists and engineers per million population	.44	.01	35
Log. GDP (1990) per capita adjusted for PPPI	.40	.01	40
Percent of the GDP in agriculture, hunting, forestry, and fishing	-.49	.00	37
Percent of labor force employed in industry	.39	.02	37
Percent of labor force employed in agriculture	-.41	.01	37

Note: Data from the Third International Mathematics and Science Study (TIMSS)

Also of importance are the many predictor variables seen in Appendix A that were not correlated cross-nationally with mean science scores. Many of these are of special interest because the TIMSS Conceptual Framework (Martin & Kelly, 1996, Figure 5.4, p. 5-8) identifies a variety of policy-relevant student, teacher, curriculum, instruction, and educational system variables that are presumed to be linked to variation in mean student achievement scores across nations. Our analyzes test whether these presumed relationships are observed cross-nationally in TIMSS data. Listed below are a number of general categories of predictor variables shown in the TIMSS Conceptual Framework to be related to variation in science scores cross nationally, followed by a relevant example of a particular nationally-aggregated variable provided by TIMSS that was not related cross-nationally to differences in mean science scores:

1. Student background: Percentage of foreign born students
2. Student activities: Percentage of students who took extra lessons in science outside school
3. Student attitudes: Percentage of students who strongly agree that to do well in science you need lots of hard work studying at home
4. Teacher qualifications: Years of teaching experience
5. Teacher status: Percentage of full-time teachers
6. Teacher social organization: Number of times teachers meet with other teachers to plan curriculum and approaches to teaching
7. Teacher pedagogic beliefs: Importance for students to memorize formulas and procedures
8. Intended curriculum: (a) Breadth or (b) Repetition
9. Instructional activities: Percentage of teaching time based on textbook
10. System characteristics: Student age at start of compulsory schooling

These 10 variables merely illustrate the many predictor variables of policy relevance that are not associated with science achievement scores cross-nationally. Though too numerous to list and discuss here, all predictor variables analyzed are found in Appendix A.

The Multivariate Model for Science Achievement

Though the bivariate relationships of each of the predictor variables with national mean scores as shown in Appendix A are of considerable interest individually, of greater interest and importance are those included in the best multivariate prediction models. Therefore, multivariate analysis were used to quantify the simultaneous association of predictor variables with science achievement and to identify the predictor variables with the largest effects. The best multivariate model[4] (as generated by the OLS method) is shown in the bottom panel of Figure 4. As seen, four variables (two based on the questionnaire responses of Principals describing their schools, and two based on questionnaire responses of students) accounted for 85% of the total variability observed among nations in mean science achievement[6]. These variables were:

1. School provides any special science enrichment activities (referred to below as the "science enrichment" variable),
2. Capacity of school to provide instruction is affected "a lot" by a shortage or inadequacy of science laboratory equipment and materials (referred to below as the "laboratory equipment shortage" variable),
3. Hours students work on paid job per week (referred to below as the "student paid job" variable), and
4. National percentage of students who strongly agreed that to do well in science they need good luck (referred to below as the "need luck in science" variable).

The power of this model to predict national mean science scores with little error is illustrated in Figure 10 by the close relationship between the observed national mean science scores and the national mean science scores predicted by this four-variable model. Note also that the six nations fall close to the regression line, with top-scoring Singapore having the very highest predicted national score.

When this four-variable prediction model was recomputed with independent national samples at the 7th grade, it accounted for 80 percent of the national-level variability in mean math achievement scores. Thus, the best cross-national multivariate model at grade 8 for predicting science achievement was replicated with independent national samples of students at grade 7.

---

[6] When this four-variable prediction model was recomputed with independent national samples at the 7th grade, it accounted for 80 percent of the national-level variability in mean math achievement scores. The OLS and HLM methods produced virtually identical estimates (about 70% for the sample of 37 nations in the HLM analysis) of national-level differences accounted for by the four-variable science model shown in Figure 4. Thus, the powerful four-variable science model, as shown in Figure 4, is very stable across national samples (7th and 8th grades) and analytic methods (OLS and HLM).

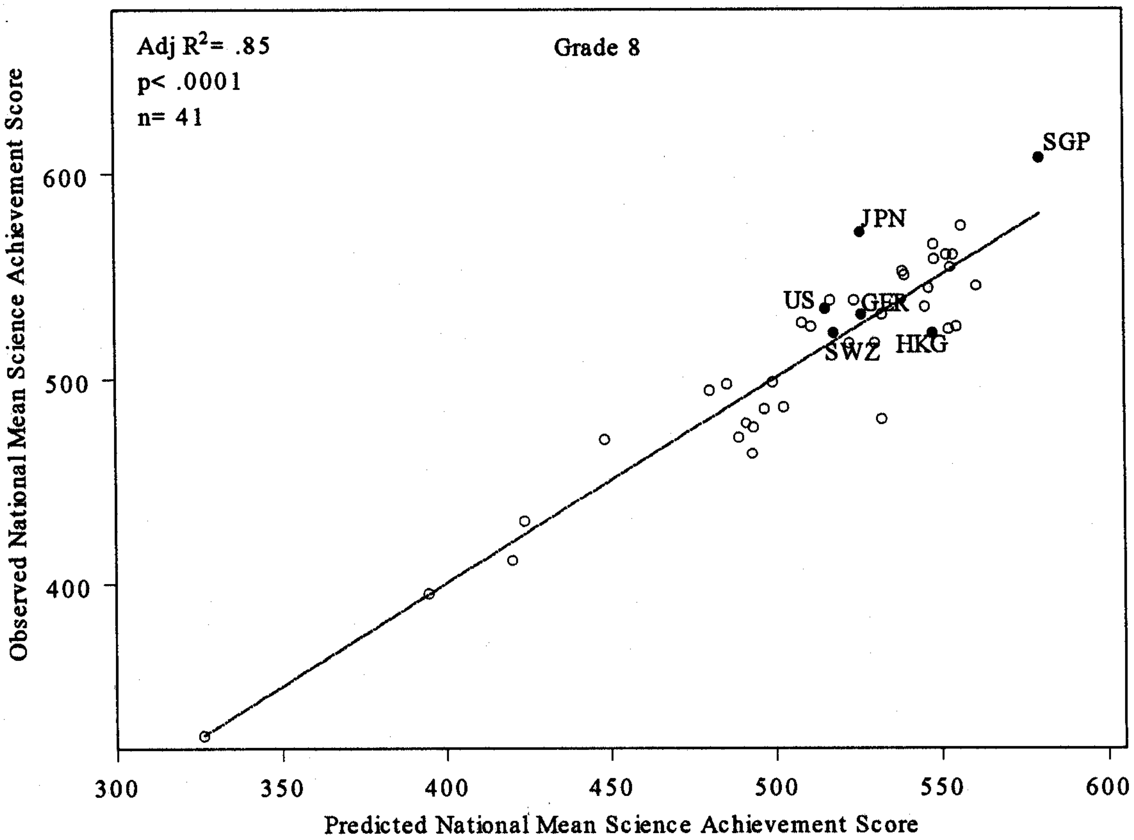


Figure 10. Observed national mean science achievement scores as a function predicted mean science scores for 41 nations. Predictions are based on the best multivariate model using four predictor variables. Six observations are highlighted: Germany (GER), Hong Kong (HKG), Japan (JPN), Singapore (SGP), Switzerland (SWZ), and the United States of America (US). Model adjusted R^2 = .85. Data Source: TIMSS.

In interpreting the results of these analyses at the national level, it is important to recognize that they are likely to be quite different from results of analyses that could be conducted at the student level. Therefore, generalizations between levels should not be made without clear evidence of their validity. To do so without such evidence is to commit the "ecological fallacy."

Multilevel Analysis for Science

In the HLM science model, the sample reduction after listwise deletion of cases with missing data was notable. Approximately 33% of the total sample of 147,505 students, 25% of the total sample of 5,879 schools, and 12% of the total sample of 41 nations were dropped for this analysis.

The HLM and OLS methods of analysis produced estimates of variance explained at the national level (approximately 71%) by the four-variable science model shown in Figure 4. The regression coefficients for each national-level predictor in the HLM model were also similar to those from the OLS method. Although not estimated using OLS, the estimate of variance explained within nations by the group-centered student-level versions of the predictor variables in the HLM model was much smaller (only 1.1%) than at the national level. Thus, the powerful four-variable national-level model was replicated using HLM, although the predictive power of the four-variable model within nations was much less.

Supplementary Analyses of Science Predictor Variables

Supplementary analyses were performed to understand better each of the four predictor variables that emerged as main effects in the best model of national achievement in science as shown in Figure 4. A policy implication is drawn from each variable as a predictor of national-level science achievement.

Additional Analyses of the Predictor Variable "Science Enrichment"

As shown in Figures 4 and 11, the variable "science enrichment" was found to be associated directly with mean science achievement cross-nationally. We next performed within-nation analyses at the student level separately for each of the 39 nations available for this analysis. As might be expected, results of these analyses showed that variable "science enrichment" was also directly related to science achievement scores within each nation. Thus, the provision of science

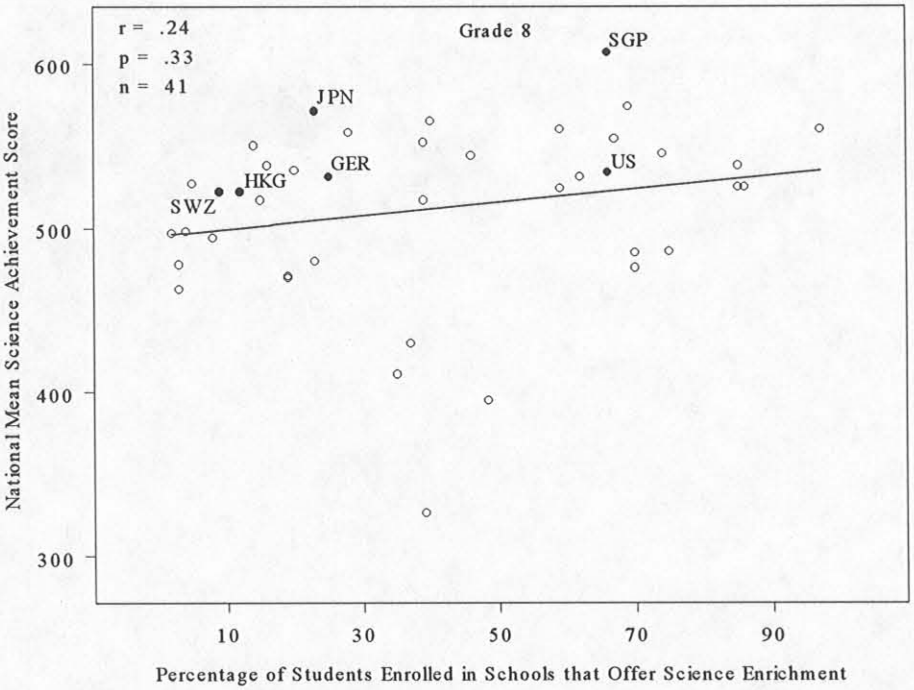


Figure 11. National mean science achievement scores as a function of the national percent of students who are enrolled in a school that provides any special science enrichment activities (based on imputed national data). Source: TIMSS.

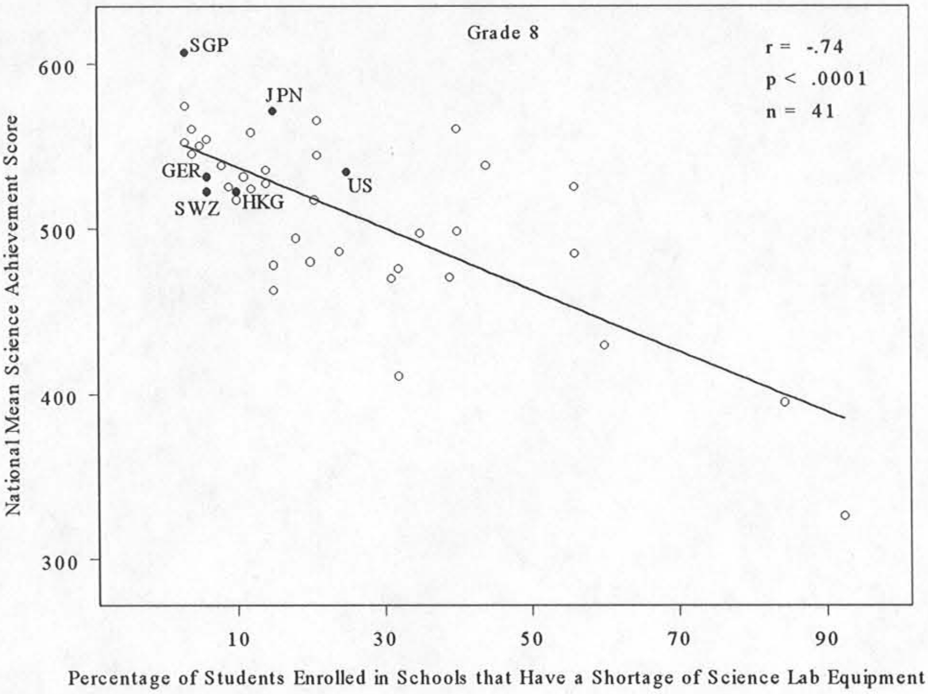


Figure 12. National mean science achievement scores as a function of the national percent of students who are enrolled in a school whose capacity to provide instruction is affected “a lot” by a shortage or inadequacy of science laboratory equipment and materials (based on imputed national data). Source: TIMSS.

enrichment activities by schools may strengthen student achievement both cross-nationally and within each nation.

Further analyses were performed to identify which other predicator variables correlated most highly cross-nationally with the variable "science enrichment." The three variables that correlated most highly with "science enrichment" were:

1. School provides any math enrichment activities (correlation with "science enrichment" = .91)
2. School provides remedial teaching in science (correlation with "science enrichment" = .61)
3. Average hours per week teachers meet with students outside of classroom time (correlation with "science enrichment" = .55)

Thus, it appears that the variable "science enrichment", and the three variables most strongly related to it, can be characterized as instructional practices that might be expected to contribute to student achievement.

In view of these consistent direct relationships with science achievement, both within and across nations, the most plausible implication of the variable "science enrichment" for educational policy is that:

- Schools should be enabled, and required, to offer science enrichment activities.

Additional Analyses of the Predictor Variable "Laboratory Equipment Shortage"

As seen in Figures 4 and 12, the variable "laboratory equipment shortage" was found to be associated inversely with mean science achievement cross-nationally. After computing within-nation associations at the student level for each of the 38 nations available for this analysis, the variable "laboratory equipment shortage" was also found to be related inversely to science achievement scores within each nation. Thus, a shortage of laboratory equipment may detract from student achievement both cross-nationally and within each nation.

In view of these consistent direct relationships with science achievement, both within and across nations, the most plausible implication of the variable "laboratory equipment shortage" for educational policy is that:

- Schools should be enabled, and required, to offer science instruction supported by adequate laboratory equipment and supplies.

Additional Analyses of the Predictor Variable "Student Paid Job"

The variable "student paid job" was similarly found to be associated inversely with mean science achievement cross-nationally (see Figures 4 and 13). After computing within-nation associations at the student level for each of the 40 nations available for this analysis, the variable "student paid job" was also related inversely to science achievement scores in all but a few nations. Thus, the extent to which parents need (or allow) their 13-year old children to work at a paid job to earn money seems to increase the time and energy that students expend on a non-academic activity, that in turn impairs their ability to perform well in science.

In view of these consistent inverse relationships between the variable "student paid job" and science achievement (both within and across nations), the implication for social policy is that:

- Family economic status should be sufficiently enhanced so that 13-year old students do not need to work at a job to earn money. Alternatively, strengthen child labor laws prohibiting employment of 13-year old adolescents.

Additional Analyses of the Predictor Variable "Need Luck in Science"

The variable "need luck in science" was also found to be associated inversely with science achievement cross-nationally (see Figures 4 and 14). After computing within-nation associations at the student level for each nation, the variable "need luck in science" was also related inversely to science achievement scores in all but a few nations. Thus, to the extent that students attribute doing well in science to luck, they seem to do poorly possibly because this detracts from their making the effort that is required to do well.

In view of these consistent inverse relationships of the variable "need luck in science" with science achievement (both within and across nations), the implication for families and educational policy is that:

- Children should be inculcated in the concept that learning in school requires hard work instead of good luck, and a milieu should be created in which the work ethic is modeled and successful.

Science Prediction Model Profile for Six Nations

Based on the results of the best multivariate model for science, a four-variable profile for predicting national average science scores is presented in Figure 15 for each of three nations from the East and three from the West. As seen in Figure 11, one predictor variable (science

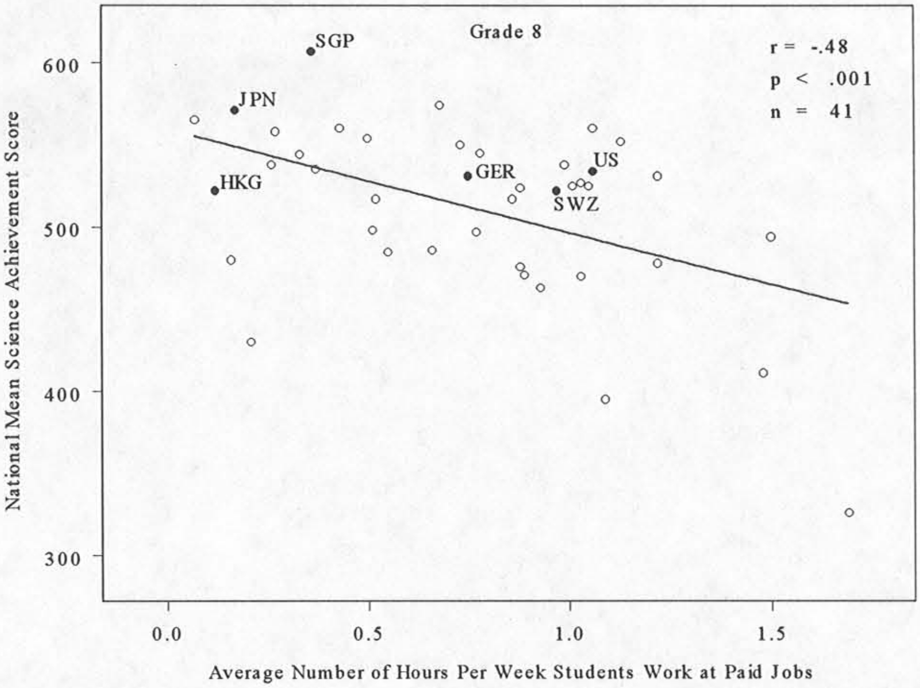


Figure 13. National mean science achievement scores as a function of the national mean number of hours per week students work at paid jobs (based on imputed national data). Source: TIMSS.

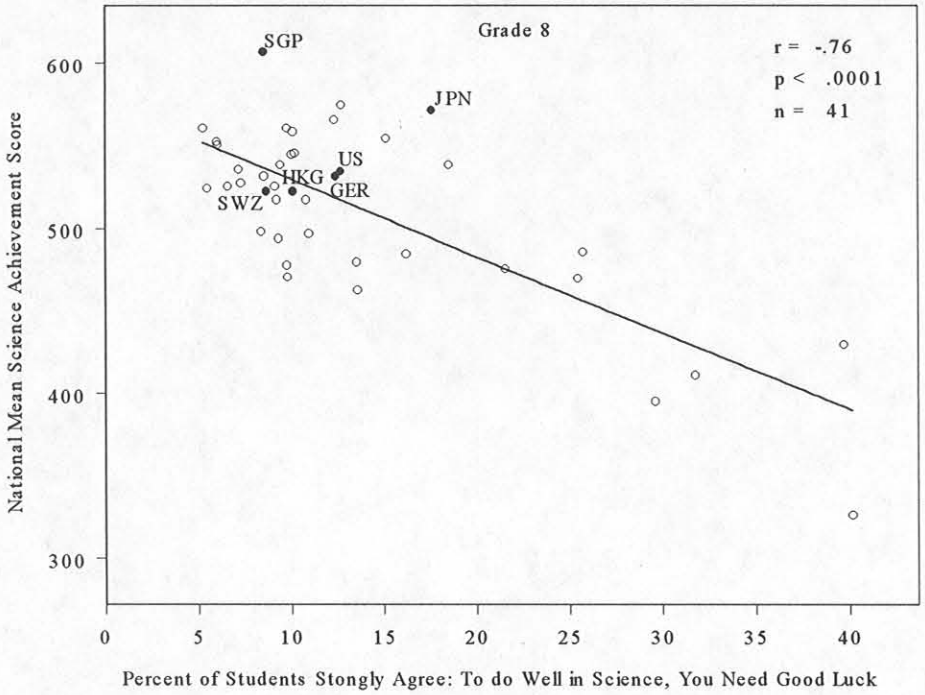
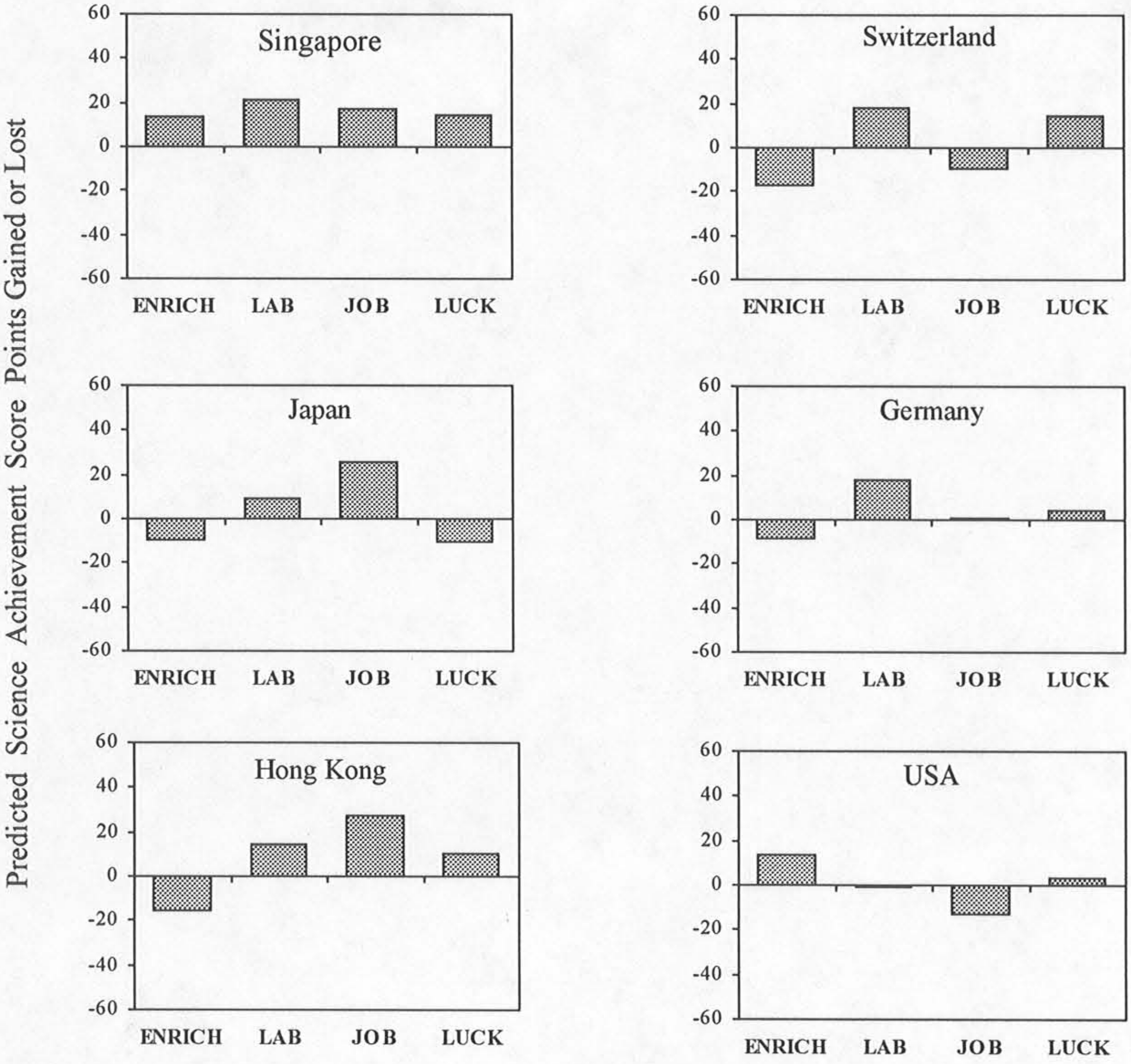


Figure 14. National mean science achievement scores as a function of the national percent of students who strongly agree that to do well in science, you need good luck (based on imputed national data). Source: TIMSS.



ENRICH: Schools provide science enrichment activities (high enrichment gains points).

LAB: Principals report a shortage of science laboratory equipment (low shortage gains points).

JOB: Number of hours students work at paid jobs per week (low hours gains points).

LUCK: Students believe they need good luck to do well in science (low belief in luck gains points)

Figure15. Predicted science achievement points gained (+) or lost (-) as a function of each of four predictor variables included in the best multivariate model for science. Points gained, or lost, are measured from the international mean science achievement score of 512. Data Source: TIMSS.

enrichment activities) was directly related to science achievement; as seen in Figures 12, 13, and 14, the other three predictor variables (laboratory equipment shortage, student paid job, need luck in science) were inversely related to science achievement cross-nationally. Accordingly, Singapore gained points on all four variables because its schools offered science enrichment activities and did not experience shortages of science laboratory equipment and supplies, and also because its students tended not to work at jobs for pay and did not believe that luck was needed to do well in science.

Perhaps the most interesting comparisons that can be made in Figure 15 are between the three nations from the East and the three nations from the West in terms of the degree to which students from each nation work at a job for pay. As seen in Figure 13, the level of national science achievement decreased as the amount of time students worked at a paid job increased (an inverse relationship). In the three Eastern nations (Singapore, Japan, and Hong Kong), the degree to which students worked for pay was well below the international average as reflected in the prediction points gained by the "Job" variable. In contrast, the degree to which students worked for pay in two Western nations (Switzerland and USA) was above average, as reflected in the prediction points lost because of the "Job" variable. As stated above, we interpret increasing degrees of student time working for pay as a distraction of time and energy that impairs their ability to perform well in science.

Comparison of Results for Math and Science

The best multivariate prediction models for math and science achievement at the 8th grade can be compared directly in Figure 4. The two models have only one predictor variable in common ("need luck to do well in math," or science). This suggests that national contexts for yielding cross-national differences in mean math achievement and mean science achievement are sufficiently different to justify separate analyses of each. For both subject matters, however, our interpretation of the multivariate results show that both educational variables and non-educational variables are major sources of national differences in mean achievement scores.

With respect to educational variables, national science achievement scores are determined strongly by instructional resource variables (specifically, enrichment activities and shortage of laboratory equipment), whereas national math achievement scores are (in our interpretation) determined strongly by national standards for teaching and learning. It seems reasonable that shortages of laboratory (and other) facilities was not a main factor in predicting national mean math scores because such facilities are not needed to teach math. In contrast, laboratory shortage was a main factor in predicting national mean science scores--a reasonable finding since laboratory facilities are a necessary resource for high quality science teaching.

Both the math and science models included variables that some might classify as "cultural." National variation of 8th grade students working at a paid job was inversely related to national mean science scores, whereas student belief in the emphasis their mothers placed on doing well in sports was likewise inversely related to national mean math scores. Both predictor variables suggest cultural differences in student activities out-of-school, activities that appear to detract from a focus on academic performance. Though this seems plausible, it is not clear why the variable "mother sports" emerged as a strong predictor in the math model but not the science model, and why the variable "student paid job" emerged in the science model, but not in the math model.

A student "attitude" ("need luck to do well in math," or science) was the common predictor variable in both the math and science multivariate models. This certainly indicates a detached perspective by students, and may reflect low motivation to learn math and science that, for most, are difficult subjects to master.

Assessment of Methods and Results: Strengths and Limitations

Among several possible approaches to studying TIMSS data, we chose to capitalize on its cross-national dimension in order to explain the sources of national differences in math and science achievement; that is, to identify the national-level variables that are the strongest predictors of variation among nations in average achievement scores. Within the strengths and limitations of TIMSS data (as augmented by national level variables from other sources), we have accomplished the first extensive national-level analysis to be reported. Our comparative education research has had several distinctive features:

- It has focused on national-level predictor variables, including student, classroom, and school variables aggregated at the national level.
- It has focused on national-level outcome variables; specifically, comprehensive (i.e., full-scale) measures of student achievement in math and science aggregated at the national level.
- It has focused on studying the cross-national relationships among these national-level predictor and outcome variables.

- It has employed sophisticated multilevel statistical analysis methods to identify key predictors of national mean achievement in math and in science.

Our analyses have shown that educational factors represent but one major source of differences among nations in academic achievement. Our interpretation of these findings is that variation in national average math achievement scores is related to the rigor of national standards for teaching and learning, whereas variation in national average science achievement scores is related to the adequacy of science laboratories and to the provision of enrichment learning opportunities. Other than these educational factors, national differences in average math and science achievement scores do not seem to be related to national differences in the breadth and repetition of curricular content, particular instructional methods, or teacher qualifications. Such variables may be important determinants of achievement differences among students within nations, but do not account for average achievement differences across nations.

By contrast, national mean math and science scores are more strongly related (generally) to student variables (i.e., background, behavior, and attitudes) than they are to instructional variables (i.e., instructional methods, homework assigned, and use of student assessments). In fact, many instructional variables that are commonly regarded as important for enhancing student achievement (e.g., frequency with which teachers meet to plan curriculum and instrctional methods) were not found to be related to cross-national achievement differences. This can be viewed as disappointing from an education policy-making perspective because the results provide little direct information about actions that can be taken to improve student achievement. Two possibilities might explain these findings:

1. Variation among nations in a wide variety of educational predictor variables may not be major sources of national differences in academic achievement, as the results of our analyses generally suggest, or

2. TIMSS data, and/or the methods we used to analyze the cross-national relationships between educational variables and student achievement, may be inadequate for detecting genuine differences among national educational systems and processes that have major influences on national differences in educational achievement.

Although we believe the first explanation to be substantially valid, there are several aspects of the design of TIMSS and our analyses that may not be insensitive to subtle cross-national educational effects. If the intent is to detect the subtle effects of national differences in educational inputs <u>during a particular school year</u> (e.g., the 8th grade) on national differences in student achievement, then the design of TIMSS is far from optimal for this purpose. As stated by the U.S. National Center for Education Statistics (NCES, 2000):

> Ideally one would like to link the attributes of teachers and teachers' instructional practices to the demonstrated achievement of the students they teach and, in this way, identify effective teachers and effective teaching practice. Such a linking is possible within the TIMSS data, but cross-sectional designs of the kind that characterize TIMSS (and IEA studies in general) are not well-suited to this propose. Students enter the eighth grade with knowledge, beliefs, and orientations accumulated over 7 years of schooling and some 13 to 14 years of family life. What teachers do within the pace of a school year is unlikely to radically alter the achievement level of the class as a whole and so create a sizable correlation between teacher instructional practices and student achievement at the classroom level. The best hope to demonstrate the relationship between teacher' instructional practices and student achievement is to look a t the relationship to growth in achievement over the year, rather than absolute levels of achievement. Recognizing this, the original design of TIMSS was one that required a pre- and posttest to measure this growth. Unfortunately, most of the participating nations were unable to support both a pre- and posttest, so the study reverted to a simple cross-sectional single testing design (p. 91).

In addition to the absence of pre- and posttest design, Schmidt et al. (1999) identified another difficulty. They observed that comprehensive (full-scale) test scores (such as in math and science) will not be as sensitive to particular instructional or curricular inputs as a more focused test, such as subscale scores in algebra, geometry, etc. If so, our analysis using full-scale achievement scores may have precluded the detection of subtle educational influences.

As noted by NCES (2000) above, "students enter the eighth grade with knowledge, beliefs, and orientations accumulated over 7 years of schooling and some 13 to 14 years of family life." It is therefore reasonable that these cumulative student characteristics might correlate highly with the full-scale achievement scores in math and science that represent cumulative learning during a students entire past years of schooling. This should be contrasted with studying particular instructional events during a particular year, the effects of which could best be measured with a pre- posttest design based on subscale scores focused on these particular instructional events.

Thus, it is possible that strong cross-national education effects might have been detected if it were not for limitations inherent in TIMSS with its cross-sectional instead of a pre- posttest design, and if particular subscale scores (rather than full-scale scores) were used. With respect to the latter point, however, it should be recognized that the full-scale and subscale scores, in both math and science, correlate from 0.94 to 0.98 at the cross-national level at which our analyses were made. Therefore, finer grain analyses with subscale scores would probably not lead to substantially different findings about the key predictors of national differences in achievement. Thus, it may well be true that, contrary to common wisdom, the predominant sources of national differences in educational achievement are not educational variables, and that this would be demonstrated with a pre- and posttest design based on subscale scores.

If the interest is in the effects of a wide range of student, educational, and broad national variables on full-scale scores in math and science, then the results reported here, based on rigorous cross-national analyses using advanced statistical methods, are no doubt valid. It should be recognized that great interest has been shown by researchers, policy makers, the press, and the public in the national differences in these comprehensive (full-scale) achievement scores as shown in Figures 1 and 2. It is cross-national variation in these scores that our research has addressed--the very scores that have generated so much interest.

References

Baker, D. P., Goesling, B., & LeTendre, G. K. (in press). Socioeconomic status, school quality, and national economic development: a cross-national analysis of the "Heyneman-Loxley effect" on mathematics and science achievement. **Comparative Education Review.**

Beaton, A. E., Martin, M. O., Mullis, I. V. S., Gonzalez, E. J., Smith, T A., & Kelly, D. L. (1996). **Science Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study (TIMSS).** Chestnut Hill, MA: Center for the Study of Testing, Evaluation, and Educational Policy, Boston College.

Beaton, A. E., Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1996). **Mathematics Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study (TIMSS).** Chestnut Hill, MA: Center for the Study of Testing, Evaluation, and Educational Policy, Boston College.

Boe, E. E., Barkanic, G., Leow, C. S., May, H., Shin, S, Singleton, J. C., et al. (2001). **Correlates of national differences in mathematics and science achievement: Evidence from TIMSS** (Data Analysis Report No. 2001-DAR1). Philadelphia: University of Pennsylvania, Graduate School of Education, Center for Research and Evaluation in Social Policy.

Gonzales, P., Calsyn, C., Jocelyn, L., Mak, K., Kastberg, D., Arafeh, S., et al. (2000). **Pursuing Excellence: Comparisons of International Eighth-Grade Mathematics and Science Achievement from a U.S. Perspective, 1995 and 1999.** Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Martin, M. O., & Kelly, D. L. (Eds.) (1996). **Third International Mathematics and Science Study Technical Report: Volume I: Design and Development.** Chestnut Hill, MA: Center for the Study of Testing, Evaluation, and Educational Policy, Boston College.

Martin, M. O., Mullis, I. V. S., Gregory, K. D., Hoyle, C., & Shen, C. (2000). **Effective schools in science and mathematics: IEA's Third International Mathematics and Science Study.** Chestnut Hill, MA: TIMSS International Study Center, Boston College.

National Center for Education Statistics (2000). **Mathematics and science in the eighth grade: Findings from the Third International Mathematics and Science Study** (NCES 2000-014). Office of Educational Research and Improvement, U.S. Department of Education: Washington, DC.

National Education Goals Panel. (1991). **The national education goals report: Building a nation of learners.** Washington, DC: Author.

Neter, J., Kutner, M. H., Nachtschiem, C. J., & Wasserman, W. (1996) **Applied Linear Statistical Models** (4th ed.). Chicago: Irwin.

Peak, L. (1996). **Pursuing excellence: A study of U.S. eighth-grade mathematics and science teaching, learning, curriculum, and achievement in international context.** Washington, DC.: National Center for Education Statistics, U.S. Department of Education.

Raudenbush, S. W., & Bryk, A. S. (2002). **Hierarchical Linear Models: Applications and Data Analysis Methods**, Second Edition. Newbury Park, CA: Sage.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., & Congdon, R. (2000). HLM (Version 5) [Computer software]. Lincolnwood, IL: Scientific Software International.

Rubin, D. B. (1987). **Multiple Imputation for Nonresponse in Surveys.** New York: Wiley.

Schmidt, W. H., McKnight, Curtis C., Gogan, L. S., Jakwerth, P. M., & Houang, R. T. (1999). **Facing the consequences: Using TIMSS for a closer look at U.S. mathematics and science education.** Dordrecht, Netherlands: Kluwer Academic Publishers.

Shafer, J. L. (1997). **Analysis of Incomplete Multivariate Data.** London: Chapman & Hall.

Schafer, J.L. (2000). NORM: Multiple imputation of incomplete multivariate data under a normal model (Version 2) [Computer software]. Retrieved from http://www.stat.psu.edu/~jls/ misoftwa.html

Woessmann, L. (2001). Why students in some countries do better: International evidence on the importance of education policy. **Education Matters**, Summer, 67-74.

APPENDIX A

CROSS-NATIONAL BIVARIATE CORRELATIONS FOR MATHEMATICS AND SCIENCE SCORES

Table A-1. *Cross-National Bivariate Correlations of Predictor Variables with Mean Math and Science Achievement Scores at the 7th and 8th Grades with Nation as Unit of Analysis (N)*

Predictor Variable				Correlation			
				Math		Science	
Subcategory	Content	Coding	Statistic	7th	8th	7th	8th
<u>Student Background and Behavior</u>							
Learning Assets in the Home	Number of books	Mean	r	.32	.35	.32	.36
			p	.05	.03	.05	.02
			N	38	40	37	39
	Have calculator	Percent	r	.44	.47	.39	.44
			p	.01	.00	.02	.00
			N	38	40	37	39
	Have computer	Percent	r	.35	.35	.36	.38
			p	.03	.03	.03	.02
			N	38	40	37	39
	Have desk or table for your own use	Percent	r	.41	.48	.20	.31
			p	.01	.00	.23	.06
			N	38	40	37	39
	Have desk or table for your own use [TRIMMED]	Percent	r	.33	.46		
			p	.04	.00	--	--
			N	37	39		
	Have dictionary	Percent	r	.37	.35	.19	.23
			p	.02	.03	.25	.16
			N	38	40	37	39
	Have calculator, computer, desk, and dictionary	Percent	r	.40	.41	.40	.43
			p	.01	.01	.01	.01
			N	38	40	37	39
Student Characteristics	Age at grade level	Mean	r	-.21	-.39	-.02	-.21
			p	.19	.01	.93	.20
			N	39	41	38	40
	Age at grade level [TRIMMED]	Mean	r	-.03	.20	.26	.28
			p	.83	.20	.13	.11
			N	38	38	37	38
	Student task persistence	# items answered / # items available	r	.72	.79	.74	.68
			p	.00	.00	.00	.00
			N	39	41	38	40
	Percent boys in school	Percent	r	.18	.21	.09	.22
			p	.28	.18	.58	.18
			N	39	41	38	40
	Native or foreign born	Percent native born	r	-.14	.07	-.07	-.06
			p	.42	.68	.67	.70
			N	37	39	36	38
	Speak language of test at home	% responding "always" or "almost always"	r	.04	.16	.07	.19
			p	.82	.31	.70	.24
			N	38	40	37	39

<u>Note</u>: Data from the Third International Mathematics and Science Study (TIMSS)

[a] na = data not available; pd = poor distribution; -- = trimming not necessary

Table A-1 (continued). *Cross-National Bivariate Correlations of Predictor Variables with Mean Math and Science Achievement Scores at the 7th and 8th Grades with Nation as Unit of Analysis (N)*

Predictor Variable				Correlation			
				Math		Science	
Subcategory	Content	Coding	Statistic	7th	8th	7th	8th
Student Characteristics (continued)	Mother born in country	Percent	r	-.40	-.27	-.25	-.15
			p	.01	.10	.13	.38
			N	37	39	36	38
	Mother born in country [TRIMMED]	Percent	r	-.31			
			p	.06	--	--	--
			N	36			
	Father born in country	Percent	r	-.40	-.30	-.24	-.16
			p	.02	.07	.15	.32
			N	37	39	36	38
	Father born in country [TRIMMED]	Percent	r	-.30			
			p	.07	--	--	--
			N	36			
Student Activities Out-of-School	Spend any time on extra lessons in math/science	Percent	r	-.21	-.23	-.28	-.37
			p	.20	.14	.08	.02
			N	39	41	38	40
	Spend any time on extra lessons in math/science [TRIMMED]	Percent	r			.04	-.03
			p	--	--	.77	.86
			N			36	38
	Hours per week at paid job	Mean	r	-.66	-.51	-.52	-.36
			p	.00	.00	.00	.02
			N	38	40	37	39
	Hours per day spent watching TV or videos	Mean	r	.22	.31	.28	.39
			p	.17	.05	.09	.01
			N	39	41	38	40
	Hours per day spent playing computer games	Mean	r	-.16	-.11	-.09	.01
			p	.33	.48	.58	.94
			N	39	41	38	40
	Hours per day spent playing or talking with friends outside of school	Mean	r	-.00	.09	.12	.22
			p	1.00	.58	.48	.17
			N	39	41	38	40
	Hours per day spent doing jobs at home	Mean	r	-.59	-.58	-.40	-.44
			p	.00	.00	.01	.00
			N	39	41	38	40
	Hours per day spent playing sports	Mean	r	-.30	-.35	.00	-.10
			p	.07	.02	.98	.54
			N	39	41	38	40
	Hours per day spent reading a book for enjoyment	Mean	r	-.39	-.40	-.35	-.32
			p	.01	.01	.03	.05
			N	39	41	38	40

Note: Data from the Third International Mathematics and Science Study (TIMSS)

[a] na = data not available; pd = poor distribution; -- = trimming not necessary

Table A-1 (continued). *Cross-National Bivariate Correlations of Predictor Variables with Mean Math and Science Achievement Scores at the 7th and 8th Grades with Nation as Unit of Analysis (N)*

Predictor Variable				Correlation			
				Math		Science	
Subcategory	Content	Coding	Statistic	7th	8th	7th	8th
Student Activities Out-of-School (continued)	Hours per day spent reading a book for enjoyment [TRIMMED]	Mean	r	--	.03	--	--
			p		.84		
			N		39		
	Hours per day spent on math homework	Mean	r	-.40	-.54	-.38	-.49
			p	.01	.00	.02	.00
			N	38	40	37	39
	Hours per day spent on math homework [TRIMMED]	Mean	r	-.54	-.68	--	--
			p	.00	.00		
			N	37	39		
	Hours per day spent on science homework	Mean	r	-.39	-.54	-.27	-.41
			p	.01	.00	.11	.01
			N	38	40	37	39
	Hours per day spent on science homework [TRIMMED]	Mean	r	--	--	-.36	-.54
			p			.03	.00
			N			36	38
	Hours per day spent on math & science homework	Mean	r	-.41	-.55	-.33	-.46
			p	.01	.00	.04	.00
			N	38	40	37	39
	Hours per day spent on math & science homework [TRIMMED]	Mean	r	-.54	-.69	-.42	-.59
			p	.00	.00	.01	.00
			N	37	39	36	38
	Hours per day spent on other homework	Mean	r	-.35	-.53	-.37	-.50
			p	.03	.00	.03	.00
			N	38	40	37	39
	Hours per day spent on total homework	Mean	r	-.40	-.55	-.35	-.49
			p	.01	.00	.03	.00
			N	38	40	37	39
Student Activities In-School	Students often have tests during math/science lessons	% responding "never"	r	-.10	-.13	.02	-.10
			p	.53	.40	.91	.55
			N	39	41	38	40
	Students work in small groups during math/science lessons	% responding "never"	r	-.49	-.39	-.02	-.10
			p	.00	.01	.91	.54
			N	39	41	38	40
Student Attitudes, Beliefs and Perceptions							
Student's Beliefs	It is important to do well in math/science at school	% responding "yes" or "strongly agree"	r	-.36	-.39	-.44	-.47
			p	.02	.01	.01	.00
			N	39	41	38	40
	I usually do well in math/science	% responding "strongly agree"	r	-.59	-.60	-.57	-.45
			p	.00	.00	.00	.00
			N	39	41	38	40

Note: Data from the Third International Mathematics and Science Study (TIMSS)

[a] na = data not available; pd = poor distribution; -- = trimming not necessary

Table A-1 (continued). *Cross-National Bivariate Correlations of Predictor Variables with Mean Math and Science Achievement Scores at the 7th and 8th Grades with Nation as Unit of Analysis (N)*

Predictor Variable				Correlation			
				Math		Science	
Subcategory	Content	Coding/Source	Statistic	7th	8th	7th	8th
Student's Beliefs (continued)	It is important to do well in language at school	% responding "strongly agree"	r	-.32	-.39	-.41	-.42
			p	.05	.01	.01	.01
			N	39	41	38	40
	It is important to have time to have fun	% responding "yes" or "strongly agree"	r	.04	.18	.09	.17
			p	.83	.26	.58	.30
			N	39	41	38	40
	It is important to be good at sports	% responding "yes" or "strongly agree"	r	-.59	-.49	-.49	-.44
			p	.00	.00	.00	.00
			N	39	41	38	40
	It is important to be in class with high achieving students	% responding "strongly agree"	r	-.55	-.58	-.57	-.60
			p	.00	.00	.00	.00
			N	38	40	37	39
	To do well in math/science at school you need lots of natural talent/ability	% responding "strongly agree"	r	-.52	-.49	-.52	-.53
			p	.00	.00	.00	.00
			N	38	40	37	39
	To do well in math/science at school you need good luck	% responding "strongly agree"	r	-.69	-.72	-.65	-.67
			p	.00	.00	.00	.00
			N	38	40	37	39
	To do well in math/science at school you need to do hard work studying at home	% responding "strongly agree"	r	-.22	-.26	-.28	-.27
			p	.19	.11	.10	.10
			N	38	40	37	39
	To do well in math/science at school you need to memorize the textbook or notes	% responding "strongly agree"	r	-.36	-.43	-.25	-.29
			p	.03	.01	.14	.08
			N	38	40	36	38
	I like math/science	% responding "like it a lot"	r	-.47	-.54	-.49	-.48
			p	.00	.00	.00	.00
			N	39	41	38	40
	I enjoy learning math/science	% responding "strongly agree"	r	-.73	-.76	-.78	-.74
			p	.00	.00	.00	.00
			N	38	40	38	40
	Math/science is boring	% responding "strongly agree"	r	-.36	-.26	-.10	-.09
			p	.03	.11	.54	.60
			N	38	40	38	40
	Math/science is easy	% responding "strongly agree"	r	-.75	-.76	-.66	-.68
			p	.00	.00	.00	.00
			N	38	40	38	40
	Math/science is important to everyone's life	% responding "strongly agree"	r	-.59	-.59	-.67	-.69
			p	.00	.00	.00	.00
			N	38	40	38	40
	I would like a job that involved using math/science	% responding "strongly agree"	r	-.76	-.78	-.67	-.68
			p	.00	.00	.00	.00
			N	38	40	38	40

Note: Data from the Third International Mathematics and Science Study (TIMSS)

[a] na = data not available; pd = poor distribution; -- = trimming not necessary

Table A-1 (continued). *Cross-National Bivariate Correlations of Predictor Variables with Mean Math and Science Achievement Scores at the 7th and 8th Grades with Nation as Unit of Analysis (N)*

Predictor Variable				Correlation			
				Math		Science	
Subcategory	**Content**	**Coding/Source**	**Statistic**	**7th**	**8th**	**7th**	**8th**
	I need to do well in math/science to get the job I want	% responding "strongly agree"	r	-.72	-.69	-.56	-.54
			p	.00	.00	.00	.00
			N	39	41	38	40
Student's Beliefs (continued)	I need to do well in math/science to please myself	% responding "strongly agree"	r	-.55	-.59	-.61	-.61
			p	.00	.00	.00	.00
			N	39	41	22	23
	I need to do well in math/science to get into the secondary school or university I prefer	% responding "strongly agree"	r	-.62	-.53	-.42	-.40
			p	.00	.00	.01	.01
			N	39	41	38	40
	In my math class students often neglect their schoolwork	% responding "strongly agree"	r	-.77	-.61	na[a]	na[a]
			p	.00	.00		
			N	38	40		
Mother's Beliefs (Student report)	Mother thinks it is important to do well in math/science	% responding "yes" or "strongly agree"	r	-.23	.03	-.45	-.48
			p	.16	.86	.01	.00
			N	38	40	37	38
	Mother thinks it is important to do well in language	% responding "strongly agree"	r	-.23	-.25	-.37	-.34
			p	.16	.12	.02	.04
			N	38	39	37	38
	Mother thinks it is important to be good at sports	% responding "yes" or "strongly agree"	r	-.80	-.76	-.69	-.67
			p	.00	.00	.00	.00
			N	38	39	37	38
	Mother thinks it is important to have time to have fun	% responding "yes" or "strongly agree"	r	-.25	-.20	-.14	-.16
			p	.13	.23	.42	.34
			N	38	39	37	38
	Mother thinks it is important to be in class with high achieving students	% responding "strongly agree"	r	-.49	-.44	-.57	-.52
			p	.00	.01	.00	.00
			N	37	38	36	37
Friends' Beliefs (Student report)	My friends think it is important to do well in math/science	% responding "yes" or "strongly agree"	r	-.48	-.44	-.55	-.53
			p	.00	.00	.00	.00
			N	39	41	38	40
	My friends think it is important to do well in language	% responding "strongly agree"	r	-.41	-.40	-.55	-.47
			p	.01	.01	.00	.00
			N	39	41	38	40
	My friends think it is important to do well in language [TRIMMED]	% responding "strongly agree"	r	--	--	--	-.68
			p				.00
			N				39
	My friends think it is important to be good at sports	% responding "yes" or "strongly agree"	r	-.66	-.62	-.54	-.54
			p	.00	.00	.00	.00
			N	39	41	38	40

Note: Data from the Third International Mathematics and Science Study (TIMSS)

[a] na = data not available; pd = poor distribution; -- = trimming not necessary

Table A-1 (continued). *Cross-National Bivariate Correlations of Predictor Variables with Mean Math and Science Achievement Scores at the 7th and 8th Grades with Nation as Unit of Analysis (N)*

Predictor Variable				Correlation			
				Math		Science	
Subcategory	Content	Coding	Statistic	7th	8th	7th	8th
Friends' Beliefs (Student report) (continued)	My friends think it is important to have time to have fun	% responding "yes" or "strongly agree"	r	.06	.20	.11	.19
			p	.73	.20	.50	.24
			N	39	41	38	40
	My friends think it is important to be in class with high achieving students	% responding "strongly agree"	r	-.54	-.57	-.61	-.60
			p	.00	.00	.00	.00
			N	38	40	37	39
Student Safety and Discipline (Student report)	I skipped a class last month	% responding "no" or "never"	r	-.55	-.52	-.54	-.48
			p	.00	.00	.00	.00
			N	35	37	34	36
	Something of mine was stolen last month	% responding "no" or "never"	r	-.28	-.25	-.12	-.17
			p	.10	.14	.50	.32
			N	35	37	34	36
	I was threatened/hurt by another student last month	% responding "no" or "never"	r	-.36	-.29	-.29	-.24
			p	.03	.09	.10	.16
			N	35	37	34	36
	Some of my friends skipped classes last month	% responding "no" or "never"	r	-.50	-.43	-.53	-.45
			p	.00	.01	.00	.01
			N	35	37	34	36
	Some of my friends had things stolen last month	% responding "no" or "never"	r	-.02	.07	.06	.08
			p	.93	.68	.73	.66
			N	35	37	34	36
	Some of my friends were threatened/hurt by other students last month	% responding "no" or "never"	r	-.43	-.35	-.49	-.42
			p	.01	.03	.00	.01
			N	35	37	34	36
<u>Instructional Factors</u>							
Teacher Qualifications	Part-time teachers	Percent	r	-.16	-.14	.06	.05
			p	.35	.39	.75	.76
			N	36	39	36	38
	Part-time teachers [TRIMMED]	Percent	r	-.37	-.26		
			p	.03	.11	--	--
			N	35	38		
	Number of years teaching	Mean	r	.12	.20	.07	.23
			p	.48	.22	.70	.16
			N	36	39	36	39
	Number of years of post-secondary training required	Whole Years	r	.22	.24	.14	.18
			p	.18	.13	.40	.27
			N	39	41	38	40
	Teaching practice required	Yes/No	r	.12	.10	.10	.08
			p	.48	.55	.54	.63
			N	39	41	38	40

<u>Note</u>: Data from the Third International Mathematics and Science Study (TIMSS)

[a] na = data not available; pd = poor distribution; -- = trimming not necessary

Table A-1 (continued). *Cross-National Bivariate Correlations of Predictor Variables with Mean Math and Science Achievement Scores at the 7th and 8th Grades with Nation as Unit of Analysis (N)*

Predictor Variable				Correlation			
				Math		Science	
Subcategory	Content	Coding	Statistic	7th	8th	7th	8th
Teacher Qualifications (continued)	Teacher examination required	Yes/No	r	.26	.21	.27	.24
			p	.10	.19	.10	.14
			N	39	41	38	40
	Teacher age	Mean	r	-.03	.11	.04	.18
			p	.88	.49	.82	.27
			N	36	39	36	39
	Female teachers	Percent	r	-.08	.07	-.08	.03
			p	.65	.68	.62	.84
			N	36	39	36	39
	Teachers with MsEd or PhD	Percent	r		.23	.32	.29
			p	pd	.19	.08	.10
			N		34	31	34
Teacher Beliefs	If students have difficulty, an effective approach is... more practice by themselves	% responding "strongly agree"	r	-.08	-.22		
			p	.63	.19	na[a]	na[a]
			N	34	37		
	Some students have a natural talent for math/science and others do not	% responding "strongly agree"	r	-.18	-.22	.01	-.05
			p	.29	.18	.96	.75
			N	35	38	34	37
	Percent of students in the top third nationally (high achievement levels)	Mean	r	.07	.04	.03	.12
			p	.70	.81	.89	.50
			N	31	34	30	34
	Percent of students in the middle third nationally (middle achievement levels)	Mean	r	.05	.11	.11	.26
			p	.79	.54	.55	.14
			N	31	34	30	34
	Percent of students in the bottom third nationally (low achievement levels)	Mean	r	-.12	-.15	-.14	-.35
			p	.52	.39	.47	.04
			N	31	34	30	34
	Teaching is limited by different academic abilities	% responding "a great deal"	r	-.16	-.15	-.19	-.16
			p	.34	.37	.27	.33
			N	36	39	36	39
	Teaching is limited by range of student backgrounds	% responding "a great deal"	r	-.16	-.41	-.21	-.45
			p	.35	.01	.23	.00
			N	35	38	35	38
	Teaching is limited by uninterested students	% responding "a great deal"	r	-.28	-.26	-.30	-.30
			p	.10	.12	.08	.06
			N	36	39	36	39
	Teaching is limited by high student/faculty ratio	% responding "a great deal"	r	-.16	-.18	-.32	-.32
			p	.35	.28	.05	.05
			N	36	39	36	39
	Teaching is limited by low morale among faculty	% responding "a great deal"	r	-.16	-.34	-.19	-.35
			p	.37	.04	.28	.03
			N	35	38	35	38

Note: Data from the Third International Mathematics and Science Study (TIMSS)

[a] na = data not available; pd = poor distribution; -- = trimming not necessary

Table A-1 (continued). *Cross-National Bivariate Correlations of Predictor Variables with Mean Math and Science Achievement Scores at the 7th and 8th Grades with Nation as Unit of Analysis (N)*

Predictor Variable				Correlation			
				Math		Science	
Subcategory	Content	Coding	Statistic	7th	8th	7th	8th
Teacher Beliefs (continued)	It is important to remember formulas and procedures	% responding "very important"	r	-.05	-.23	-.03	-.21
			p	.76	.17	.87	.22
			N	35	38	35	38
	Social status rank of teaching profession	Mean	r	.29	.40	.36	.40
			p	.18	.04	.09	.04
			N	23	26	23	26
	Have influence on subject matter	% responding "a lot"	r	-.08	-.08	.10	.06
			p	.66	.64	.58	.71
			N	35	37	35	37
	Have influence on specific text used	% responding "a lot"	r	-.01	.01	.04	.03
			p	.94	.97	.80	.84
			N	35	37	35	37
Teacher Activities	Teachers divide class into groups for teaching math/science	% responding "sometimes" or "always"	r	-.15	-.14	-.12	-.16
			p	.41	.42	.52	.36
			N	33	36	33	36
	Hours per week outside school preparing exams	Mean	r	.14	.01	-.04	-.13
			p	.43	.95	.80	.44
			N	35	38	35	38
	Hours per week outside school reading student work	Mean	r	.16	.11	.18	.13
			p	.36	.51	.30	.43
			N	35	38	35	38
	Hours per week outside school planning lessons	Mean	r	.06	.21	.06	.15
			p	.75	.20	.72	.37
			N	35	38	35	38
	Hours per week outside school meeting with students	Mean	r	.37	.32	.19	.31
			p	.03	.05	.27	.06
			N	35	38	35	38
	Hours per week outside school meeting with parents	Mean	r	-.49	-.36	-.37	-.23
			p	.00	.03	.03	.16
			N	35	38	35	38
	Hours per week outside school on professional development	Mean	r	-.12	-.03	-.12	-.05
			p	.50	.86	.48	.78
			N	35	38	35	38
	Hours per week outside school keeping student records up-to-date	Mean	r	.09	.01	.14	.08
			p	.63	.97	.41	.65
			N	35	38	35	38
	Hours per week outside school on administrative tasks	Mean	r	.22	.35	.31	.46
			p	.21	.03	.07	.00
			N	35	38	35	38

Note: Data from the Third International Mathematics and Science Study (TIMSS)

[a] na = data not available; pd = poor distribution; -- = trimming not necessary

Table A-1 (continued). *Cross-National Bivariate Correlations of Predictor Variables with Mean Math and Science Achievement Scores at the 7th and 8th Grades with Nation as Unit of Analysis (N)*

Predictor Variable				Correlation			
				Math		Science	
Subcategory	Content	Coding	Statistic	7th	8th	7th	8th
Teacher Activities (continued)	Hours/periods scheduled for student supervision	Mean	r	.22	.01	.30	.09
			p	.24	.95	.11	.64
			N	31	33	31	33
	Hours/periods scheduled for student counseling	Mean	r	.05	-.05	.10	-.05
			p	.78	.78	.60	.76
			N	31	33	31	33
	Hours/periods scheduled for administrative duties	Mean	r	.26	.12	.30	.12
			p	.15	.50	.10	.50
			N	31	33	31	33
	Hours/periods scheduled for individual planning	Mean	r	.06	.12	.22	.19
			p	.76	.53	.24	.31
			N	30	31	30	31
	Hours/periods scheduled for co-operative planning	Mean	r	.42	-.47	-.22	-.33
			p	.03	.01	.24	.07
			N	29	30	29	30
	Hours/periods scheduled for other non-student contact	Mean	r	.11	.07	.23	.31
			p	.05	.12	.18	.06
			N	35	38	35	38
	Percent of teaching time based on text	Mean	r	.26	.26	.09	.08
			p	.13	.12	.61	.65
			N	35	38	35	38
	Number of curriculum planning meetings per year	Mean	r	-.16	-.20	.09	-.08
			p	.37	.22	.62	.62
			N	35	38	35	38
Homework Assigned	Times per week assign homework	Mean	r	-.24	-.24	-.28	-.35
			p	.16	.15	.10	.03
			N	36	39	35	38
	Minutes per week of homework assigned	Mean	r	-.19	-.21	-.18	-.21
			p	.27	.19	.30	.21
			N	36	39	35	38
Student Assessment	Use assessment to provide grades	% responding "a great deal"	r	-.16	-.24	-.14	-.12
			p	.37	.16	.45	.49
			N	34	37	34	37
	Use assessment to provide student feedback	% responding "a great deal"	r	-.28	-.20	-.15	-.09
			p	.10	.23	.39	.61
				35	38	35	38
	Use assessment to diagnose learning problem	% responding "a great deal"	r	-.17	-.13	-.16	-.20
			p	.33	.43	.36	.24
			N	35	38	35	38

Note: Data from the Third International Mathematics and Science Study (TIMSS)

[a] na = data not available; pd = poor distribution; -- = trimming not necessary

Table A-1 (continued). *Cross-National Bivariate Correlations of Predictor Variables with Mean Math and Science Achievement Scores at the 7th and 8th Grades with Nation as Unit of Analysis (N)*

Predictor Variable				Correlation			
				Math		Science	
Subcategory	Content	Coding	Statistic	7th	8th	7th	8th
Student Assessment (continued)	Weight of reasoning test	% responding "a great deal"	r	-.21	.17	-.06	-.01
			p	.22	.32	.73	.93
			N	35	38	35	38
	Weight of multiple-choice test	% responding "a great deal"	r	-.04	-.23	-.16	-.30
			p	.84	.17	.36	.07
			N	34	38	34	38
	Weight of project performance	% responding "a great deal"	r	-.24	-.07	-.11	.01
			p	.17	.66	.52	.96
			N	35	38	35	38
	Use curriculum-based exit exams for math/science	Percentage of students subject to exam	r	.25	.30	.37	.41
			p	.12	.07	.02	.01
			N	38	39	38	39
Broad Curriculum Variables	Intended math/science curriculum breadth (at grade level)	Average number of math/science topics to be taught	r	-.10	-.22	-.06	-.02
			p	.58	.20	.74	.93
			N	34	35	30	40
	Intended math/science curriculum breadth (over past three years)	Average number of math/science topics to be taught	r	-.12	-.22	-.11	-.07
			p	.52	.21	.55	.69
			N	34	35	30	40
	Intended math/science curriculum breadth (over all prior years)	Average number of math/science topics to be taught	r	-.11	-.16	-.04	-.04
			p	.52	.34	.84	.81
			N	34	35	30	40
	Intended math/science curriculum repetition (over past 2 years)	Average number of years a math/science topic to be taught	r	-.36	-.28	-.07	-.03
			p	.04	.10	.69	.87
			N	34	35	30	40
	Intended math/science curriculum repetition (over past 3 years)	Average number of years a math/science topic to be taught	r	-.28	-.34	-.07	-.07
			p	.11	.05	.70	.68
			N	34	35	30	40
	Intended math/science curriculum repetition (over all prior years)	Average number of years a math/science topic to be taught	r	-.10	-.13	.04	.02
			p	.59	.44	.83	.91
			N	34	35	30	40
<u>School Factors</u>							
School Time and Timing	Percent of students who begin the school year also finish the school year	Mean	r	.52	.50	.51	.51
			p	.00	.00	.00	.00
			N	37	39	37	39
	Percent of students who begin the school year also finish the school year [TRIMMED]	Mean	r	.44	.41		
			p	.01	.01	--	--
			N	36	38		

<u>Note</u>: Data from the Third International Mathematics and Science Study (TIMSS)

[a] na = data not available; pd = poor distribution; -- = trimming not necessary

Table A-1 (continued). *Cross-National Bivariate Correlations of Predictor Variables with Mean Math and Science Achievement Scores at the 7th and 8th Grades with Nation as Unit of Analysis (N)*

Predictor Variable				Correlation			
				Math		Science	
Subcategory	Content	Coding	Statistic	7th	8th	7th	8th
School Time and Timing (continued)	Age at start of compulsory schooling	Whole Years	r	-.06	-.01	-.12	-.09
			p	.72	.97	.52	.61
			N	34	36	33	35
	Years of formal schooling by grade level	Whole Years	r	.10	-.04	.22	.11
			p	.56	.82	.19	.50
			N	39	41	38	40
	Number of school years between grade 8 and end of secondary	Whole Years	r		.57		.55
			p	na[a]	.01	na[a]	.01
			N		20		20
	Number of instructional days in the school year	Mean	r	.32	.38	.35	.37
			p	.05	.02	.04	.02
			N	36	37	36	37
	Number of instructional hours in the school week	Mean	r	.06	-.01	-.05	-.15
			p	.73	.96	.77	.41
			N	32	34	32	34
	Number of years student stays with same teacher	Mean	r	-.13	-.02	-.12	-.01
			p	.47	.89	.51	.94
			N	35	37	35	37
	Number of boys repeating grade	Mean	r	-.48	-.60	-.53	-.61
			p	.00	.00	.00	.00
			N	36	38	36	38
	Number of girls repeating grade	Mean	r	-.41	-.56	-.46	-.59
			p	.01	.00	.00	.00
			N	36	38	36	38
	Number of girls repeating grade [TRIMMED]	Mean	r		-.43		
			p	--	.00	--	--
			N		37		
	Number of boys studying math/science	Mean	r				
			p	pd[a]	pd[a]	pd[a]	pd[a]
			N				
	Number of girls studying math/science	Mean	r				
			p	pd[a]	pd[a]	pd[a]	pd[a]
			N				
	Percent of students absent on a typical school day	Mean	r	-.28	-.31	-.14	-.16
			p	.10	.06	.42	34
			N	37	39	37	39
	Average class size (principal report)	Mean	r	.19	.19	.15	.13
			p	.28	.25	.39	.46
			N	35	37	35	37

Note: Data from the Third International Mathematics and Science Study (TIMSS)

[a] na = data not available; pd = poor distribution; -- = trimming not necessary

Table A-1 (continued). *Cross-National Bivariate Correlations of Predictor Variables with Mean Math and Science Achievement Scores at the 7th and 8th Grades with Nation as Unit of Analysis (N)*

Predictor Variable				Correlation			
				Math		Science	
Subcategory	Content	Coding	Statistic	7th	8th	7th	8th
Instructional Resources	Percent of teachers been at school for 5 or more years (principal report)	Mean	r	.04	.14	.14	.23
			p	.79	.41	.39	.16
			N	37	39	37	39
	Teachers meet regularly (principal report)	% responding "yes"	r	-.15	-.12	-.15	-.07
			p	.39	.48	.41	.67
			N	34	36	34	36
	Teachers responsible for textbooks (principal report)	% responding "yes"	r	-.10	.05	-.05	.04
			p	.58	.76	.76	.83
			N	34	36	34	36
	Remedial math provided (principal report)	% responding "yes"	r	.05	-.03	.15	.08
			p	.79	.85	.39	.62
			N	37	39	37	39
	Remedial science provided (principal report)	% responding "yes"	r	-.02	-.11	.15	.09
			p	.92	.49	.38	.58
			N	37	39	37	39
	Math/science enrichment classes provided (principal report)	% responding "yes"	r	.14	.15	.30	.31
			p	.41	.36	.07	.06
			N	37	39	37	39
	Science lab equipment shortage affects instruction (principal report)	% responding "a lot"	r	-.37	-.42	-.50	-.53
			p	.03	.01	.00	.00
			N	36	38	36	38
	Computer hardware shortage affects instruction (teacher report)	% responding "a lot"	r	-.38	-.29	-.28	-.06
			p	.02	.08	.10	.70
			N	36	39	35	39
	Computer software shortage affects instruction (teacher report)	% responding "a lot"	r	-.28	-.21	-.14	.02
			p	.10	.20	.43	.90
			N	36	39	36	39
	Instructional materials shortage affects instruction (teacher report)	% responding "a lot"	r	-.46	-.43	-.47	-.50
			p	.00	.01	.00	.00
			N	36	39	36	39
	Classroom equipment shortage affects instruction (teacher report)	% responding "a lot"	r	-.42	-.44	-.45	-.50
			p	.01	.00	.01	.00
			N	36	39	36	39
	Inadequate physical facilities affect instruction (teacher report)	% responding "a lot"	r	-.45	-.44	-.48	-.42
			p	.01	.01	.00	.01
			N	35	38	35	38
	Math/science class size (teacher report)	Mean	r	.15	.14	.05	-.02
			p	.37	.40	.77	.91
			N	36	39	36	39

Note: Data from the Third International Mathematics and Science Study (TIMSS)

[a] na = data not available; pd = poor distribution; -- = trimming not necessary

Table A-1 (continued). *Cross-National Bivariate Correlations of Predictor Variables with Mean Math and Science Achievement Scores at the 7th and 8th Grades with Nation as Unit of Analysis (N)*

Predictor Variable				Correlation			
				Math		Science	
Subcategory	Content	Coding	Statistic	7th	8th	7th	8th
Student Safety and Discipline (Principal report)	School deals with students arriving late at school	% responding "rarely"	r	-.18	-.03	-.11	-.02
			p	.33	.88	.56	.93
			N	32	34	32	34
	School deals with student absenteeism	% responding "rarely"	r	.11	.22	.13	.18
			p	.57	.21	.46	.31
			N	32	34	32	34
	School deals with students skipping class	% responding "rarely"	r	.17	.13	.21	.10
			p	.37	.45	.24	.55
			N	32	34	32	34
	School deals with students violating dress code	% responding "rarely"	r	-.20	-.10	-.14	-.08
			p	.28	.57	.45	.67
			N	31	33	31	33
	School deals with classroom disturbance among students	% responding "rarely"	r	-.23	-.22	-.31	-.27
			p	.22	.22	.08	.13
			N	32	34	32	34
	School deals with cheating among students	% responding "rarely"	r	-.16	-.11	-.12	-.07
			p	.37	.55	.53	.71
			N	32	34	32	34
	School deals with profanity among students	% responding "rarely"	r	.05	.08	.04	.10
			p	.77	.67	.81	.58
			N	32	34	32	34
	School deals with vandalism among students	% responding "rarely"	r	-.41	-.31	-.55	-.41
			p	.02	.08	.00	.02
			N	32	34	32	34
	School deals with theft among students	% responding "rarely"	r	-.07	-.11	-.16	-.17
			p	.69	.55	.39	.35
			N	32	34	32	34
	School deals with student intimidation or verbal abuse of other students	% responding "rarely"	r	-.08	-.08	-.25	-.18
			p	.65	.65	.16	.32
			N	32	34	32	34
	School deals with physical injury to other students	% responding "rarely"	r	.14	.16	-.00	.11
			p	.44	.35	.99	.54
			N	32	34	32	34
	School deals with student intimidation or verbal abuse of teachers or staff	% responding "rarely"	r	.04	.23	-.12	.15
			p	.84	.19	.52	.41
			N	32	34	32	34
	School deals with physical injury to teachers or staff	% responding "rarely"	r				
			p	pd[a]	pd[a]	pd[a]	pd[a]
			N				
	School deals with student tobacco use or possession	% responding "rarely"	r	-.46	-.46	-.35	-.47
			p	.01	.01	.06	.01
			N	30	32	30	32

Note: Data from the Third International Mathematics and Science Study (TIMSS)

[a] na = data not available; pd = poor distribution; -- = trimming not necessary

Table A-1 (continued). *Cross-National Bivariate Correlations of Predictor Variables with Mean Math and Science Achievement Scores at the 7th and 8th Grades with Nation as Unit of Analysis (N)*

Predictor Variable				Correlation			
				Math		Science	
Subcategory	Content	Coding	Statistic	7th	8th	7th	8th
Student Safety and Discipline (Principal report) (continued)	School deals with student alcohol use or possession	% responding "rarely"	r	-.08	-.14	-.13	-.15
			p	.69	.44	.52	.41
			N	29	31	29	31
	School deals with student illegal drug use or possession	% responding "rarely"	r	-.20	-.22	-.18	-.18
			p	.29	.23	.33	.34
			N	30	32	30	32
	School deals with student weapon use or possession	% responding "rarely"	r				
			p	pd[a]	pd[a]	pd[a]	pd[a]
			N				
	School deals with inappropriate sexual behavior among students	% responding "rarely"	r	-.04	-.08	-.15	-.10
			p	.83	.68	.46	.59
			N	28	30	28	30
National Educational Attainment	School age cohort in secondary schools	Percent	r	.27	.36	.19	.28
			p	.10	.02	.27	.09
			N	37	39	36	38
	School age cohort in secondary schools [TRIMMED]	Percent	r	.37	.46		
			p	.03	.00	--	--
			N	36	38		
	Tertiary students	Number of tertiary students per 100,000	r	.24	.25	.28	.27
			p	.15	.11	.09	.09
			N	39	41	38	40
	Scientists and engineers per million	Number of scientists and engineers per million	r	.39	.41	.40	.44
			p	.02	.01	.02	.01
			N	36	36	35	35
	Scientists and engineers per million [TRIMMED]	Number of scientists and engineers per million [TRIMMED]	r	.28	.32	.36	.40
			p	.10	.06	.04	.02
			N	35	35	34	34
National Demographic Variables	Population density (1994)	People per square km	r	.48	.44	.24	.25
			p	.00	.00	.14	.12
			N	39	41	38	40
	Population density [TRIMMED]	People per square km	r			.33	.24
			p	--	--	.04	.13
			N			37	38
	Urban population (1994)	Percent of total population	r	.44	.31	.25	.17
			p	.01	.05	.14	.32
			N	37	39	36	38
	Life expectancy	Years	r	.42	.40	.35	.31
			p	.01	.01	.03	.05
			N	39	41	38	40

Note: Data from the Third International Mathematics and Science Study (TIMSS)

[a] na = data not available; pd = poor distribution; -- = trimming not necessary

Table A-1 (continued). *Cross-National Bivariate Correlations of Predictor Variables with Mean Math and Science Achievement Scores at the 7th and 8th Grades with Nation as Unit of Analysis (N)*

Predictor Variable				Correlation			
				Math		Science	
Subcategory	Content	Coding	Statistic	7th	8th	7th	8th
National Economic Development and Wealth	Gross Domestic Product per capita adjusted by PPPI (1990)	U.S. Dollars	r	.48	.49	.39	.40
			p	.00	.00	.02	.01
			N	39	41	38	40
	World Bank wealth classifications (1997)	Four ordered categories	r	.44	.39	.46	.40
			p	.00	.01	.00	.01
			N	39	41	38	40
	Public expenditures on education per capita (1994)	International dollars	r	.31	.23	.37	.28
			p	.07	.17	.03	.09
			N	35	37	35	37
	Gross Domestic Product in manufacturing (1990)	Percent	r	.25	.30	.19	.27
			p	.16	.08	.30	.12
			N	33	34	32	33
	Gross Domestic Product in agriculture, hunting, forestry, fishing (1990)	Percent	r	-.51	-.41	-.62	-.49
			p	.00	.01	.00	.00
			N	36	38	35	37
	Percent Gross Domestic Product in agriculture, hunting, forestry, fishing [TRIMMED]	Percent	r	-.60	-.59		
			p	.00	.00	--	--
			N	35	36		
	Labor force in industry (1990)	Percent	r	.28	.35	.29	.39
			p	.09	.03	.09	.02
			N	36	38	35	37
	Percent of labor force in in agriculture (1990)	Percent	r	-.44	-.38	-.45	-.41
			p	.01	.02	.01	.01
			N	36	38	35	37
	Percent of labor force in in agriculture [TRIMMED]	Percent	r	-.62	-.52		
			p	.00	.00	--	--
			N	35	37		
	Percent of labor force female (1994)	Percent	r	.17	.28	.12	.26
			p	.33	.08	.47	.11
			N	37	39	36	38
	Percent of labor force female (1994) [TRIMMED]	Percent	r	.03	.20		
			p	.86	.23	--	--
			N	36	38		
Control Variables	East Asian vs other countries	1 = East Asian 0 = Other	r	.42	.37	.18	.12
			p	.01	.02	.29	.46
			N	38	39	38	39
	Country conforms to TIMSS sampling design	1 = Yes 0 = No	r	.32	.32	.14	.30
			p	.05	.04	.39	.06
			N	39	40	38	40

Note: Data from the Third International Mathematics and Science Study (TIMSS)

[a] na = data not available; pd = poor distribution; -- = trimming not necessary

APPENDIX B

TECHNICAL METHODS

APPENDIX B: TECHNICAL METHODS

Data Sources

The Third International Mathematics and Science Study (TIMSS) was conducted in over 40 nations at the seventh and eighth grade levels[7] in 1995 under the auspices of the International Association for the Evaluation of Educational Achievement (IEA). In TIMSS, this group is referred to as Population 2. Though not used in the research reported here, a subset of 24 of these nations participated in the study at the third and fourth grade levels (called Population 1)[8], and another subset of 21 nations participated in studying students in the final year of secondary school (called Population 3)[9]. In addition to the administration of internationally standardized achievement examinations in mathematics and science, extensive questionnaire data were collected from students, their mathematics and science teachers, and the principals of the schools in which data collection occurred. The design of TIMSS yielded large national probability samples of schools and students for each participating nation at each grade level. Overviews of TIMSS are found in Beaton, Mullis, et al. (1996, Appendix A) and Martin and Kelly (1996, Chapter 1) while a more detailed description, along with technical information, is found in Martin and Kelly (1996, 1997) and Martin and Mullis (1996).

In addition to questionnaire data from TIMSS, a considerable number of national-level predictor variables were obtained from other sources. These variables augmented those obtained from TIMSS questionnaires, and are therefore designated as "augmented" in the "Question Number" column of the table of variables listed in Appendix A. The source of these augmented variables is specified in the "Coding/Source" column of this table.

---

[7] More specifically, 41 nations participated at the seventh grade level and two additional nations participated at the eighth grade level.

[8] Two additional nations participated at the fourth grade level.

[9] One addition to these 21 nations participating in both Populations 2 and 3, on other nation only participated during the final year of secondary school.

Student and Nation Samples

The student samples selected for analysis here were separate national probability samples of public and private school students close to 12 years old (seventh grade in most TIMSS nations) and to 13 years old (eighth grade in most TIMSS nations). Nations with both achievement scores and questionnaire results for grade 8 were selected for analyses. Such nations are listed in Table B-1.

Variables Analyzed

National-Level Outcome Variables. The outcome (i.e., dependent) variables analyzed were the national mean full-scale TIMSS achievement scores in mathematics and science. The national means were computed from the weighted student sample using the first plausible value (Martin & Kelly, 1997).

National-Level Predictor Variables: As seen in Appendix A, a wide variety of national-level predictor variables for math and science achievement were analyzed in the categories of (a) student background and behaviors, (b) student attitudes, beliefs, and perceptions, (c) instructional factors, (d) school factors, and (e) national demographic and economic factors. The coding of each variable is described in Appendix A. The questionnaire item numbers, wording, and sources of these variables are described in Boe, et al. (2001, Appendix A).

Particular variables were selected for analysis for one or more of four reasons: (a) relevance to educational policy, practice, and/or theory, (b) inclusion in the comprehensive TIMSS conceptual framework (Martin & Kelly, 1996, Figure 5.4), (c) inclusion as univariate distributions in reports of the TIMSS International Study Center (Mullis, et al., 1997; Martin, et al., 1997; Beaton, Mullis, et al., 1996; Beaton, Martin, et al. 1996; and Mullis, et al., 1998), and (d) of special interest and importance to educators, the public, and the authors. Not included in this list of reasons are the past results of similar cross-national research because such research has not been feasible before TIMSS. With the much larger number of nations participating in TIMSS than in prior IEA surveys, cross-national analyses with nation as the unit of analysis have become feasible.

Table B-1. Sample of TIMSS Nations used for Regression Analyses in Mathematics (M) and Science (S)

Country	Grade Level	
	7th	8th
Australia	M/S	M/S
Austria	M/S	M/S
Belgium (Flemish)	M/S	M/S
Belgium (French)	M/S	M/S
Bulgaria		
Canada	M/S	M/S
Colombia	M/S	M/S
Cyprus	M/S	M/S
Czech Republic	M/S	M/S
Denmark	M/S	M/S
England	M/S	M/S
France	M/S	M/S
Germany	M/S	M/S
Greece	M/S	M/S
Hong Kong	M/S	M/S
Hungary	M/S	M/S
Iceland	M/S	M/S
Iran	M/S	M/S
Ireland	M/S	M/S
Israel		M/S
Italy		
Japan	M/S	M/S
Korea, Republic of	M/S	M/S
Kuwait		M/S
Latvia	M/S	M/S
Lithuania	M/S	M/S
Netherlands	M/S	M/S
New Zealand	M/S	M/S
Norway	M/S	M/S
Philippines	M/S	M/S
Portugal	M/S	M/S
Romania	M/S	M/S
Russian Federation	M/S	M/S
Scotland	M/S	M/S
Singapore	M/S	M/S
Slovak Republic	M/S	M/S
Slovenia	M/S	M/S
South Africa[a]	M	M
Spain	M/S	M/S
Sweden	M/S	M/S
Switzerland	M/S	M/S
Thailand	M/S	M/S
United States	M/S	M/S

Note: Questionnaire and achievement data were both required for a nation to be included in the correlational analyses in Mathematics (M) and Science (S) reported here.

[a] South Africa was not used for the Science (S) correlational analyses because it was a distinct outlier in most bivariate relationships examined.

Analysis Design

Given the exploratory nature of this research, several safeguards were implemented to ensure that the relationships observed were genuine and not due to chance or to anomalies in the data. These safeguards include methods for testing replicability of relationships, handling colinearity among predictors, dealing with missing data, and identifying outlying data points. The methods of analysis were as follows:

First, the national-level predictor variables were classified into the following five categories:

1. Student attitudes, beliefs, and perceptions
2. Student background and behaviors
3. Instructional factors (curriculum, teaching, teachers, and assessment)
4. School factors (e.g. class size, school resources, length of the school year, etc.)
5. National demographic/economic factors

Next, a Pearson product-moment correlation between each predictor variable and national mean mathematics achievement was computed. As a safeguard against multiple statistical tests and the possibility of spurious correlations, only those variables whose correlations with achievement were statistically significant ($p < .10$) for both the 7th and 8th grade samples were selected for further analysis.

The following stage of the analysis explicitly dealt with the possibility of high colinearity, or redundancy, between pairs of predictor variables. Pairs with high colinearity were defined as any two predictor variables having a Pearson correlation greater than .69. In order to reduce redundancy in the data, each pair of collinear predictors were used as independent variables in a two-predictor regression model of national mean mathematics achievement. Based on a conceptual analysis of the results of each model, a decision was made to either (a) keep both variables for use in further modeling, or (b) keep only one of them, using the following principles as guidelines:

1. If the parameter estimates for both variables decreased (from their bivariate levels) but remained statistically significant, both variables were retained.

2. If the parameter estimate of one variable decreased markedly more than the other and to a non-statistically significant level, and one variable had a considerably higher r^2 than the other, the variable with the higher r^2 was retained. Alternatively, if the r^2 for the two variables were very similar, the variable with the statistically significant parameter in the

two-predictor variable model was retained. In deciding which variable to keep, the theoretical interest/logic of each variable was also considered.

3. If both parameter estimates decreased and lost statistical significance, a decision on which variable to keep was based on the comparative theoretical interest/logic of the two variables.

Because many variables retained after this selection process had missing values for one or more nations (however, only three variables had more than three missing values), a multiple imputation method (Rubin, 1987) was used to eliminate the combined loss of observations due to missing data. This imputation of missing values was carried out using NORM 2.03 (Schafer, 2000).

The Gaussian imputation model implemented by NORM included all predictor variables retained after the variable selection stages and the national mean mathematics achievement variable. Skewed or non-normal variables were transformed to approximate normality. The ridge prior option was used to improve and accelerate convergence. Ultimately, five "plausible values" were produced for each missing value and used to produce five complete data sets. The plausible values were individually inspected to determine whether any imputed values were illogical or beyond the limits of reasonableness.

The regression modeling procedure was carried out using all five imputed data sets. For each data set, the models with the highest R^2 for one to seven variables were produced. The R^2 values for these models were inspected to identify the number of variables needed to nearly maximize the variance explained by the model (i.e., the addition of another variable would not increase R^2 more than .03). The models containing this optimal number of variables were then inspected to identify the specific variables that consistently produced the highest R^2 across all five imputed data sets. This model was identified as the best multivariate model of mean national mathematics achievement. Final R^2 and adjusted R^2 values were computed by averaging these statistics for this best model across the five data sets. Partially standardized parameter estimates were computed by multiplying the raw regression coefficients by the standard deviation of the respective predictor variable averaged across the five imputed data sets.

Further analysis of the variables comprising the best multivariate model of mathematics achievement were carried out using hierarchical linear modeling (HLM) (Raudenbush & Bryk, 2002). This allowed the simultaneous investigation of the effects of these variables at the national level and within nations. These models were based on non-imputed data and used listwise deletion of cases with missing values. Each HLM model consisted of three levels (students, schools, and nations) and was estimated using the HLM5 software package (Raudenbush, Bryk, Cheong, & Congdon, 2000). Sampling weights were used at the student level, although normalized so that the cumulative weight of each nation was the equal. Each national-level variable in the model was also included as a group-centered variable at either the student or school level coinciding with the original level of measurement. Estimates of variance explained by the model within nations and at the national level were produced by comparing the variance component estimates from the full model to those from an unconditional (no predictors) model. Parameter estimates for variables at the national level and for continuous variables at the school and student levels were rescaled into partially standardized coefficients by multiplying by the standard deviation at the appropriate level of the respective predictor variable.