DOCUMENT RESUME

ED 475 357 TM 034 833

AUTHOR	Wang, Jianjun
TITLE	An Analysis of Relationships between Mathematics and Science Achievement in TIMSS and TIMSS-R.
PUB DATE	2003-04-00
NOTE	19p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, April 21-25, 2003).
PUB TYPE	Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	Articulation (Education); *Correlation; *Grade 8; *Junior High School Students; Junior High Schools; *Mathematics Achievement; *Science Achievement; *Scores

ABSTRACT

Articulation of mathematics and science education is advocated in the official documents of several professional organizations. To assess the benefit of curriculum integration, a national indicator needs to be developed from a correlation study of performance scores between the two subjects. In this study, a correlation analysis was conducted at the eighth grade level using international databases from the Third International Mathematics and Science Study (TIMSS) and its repetition in 1999 (TIMSS-R). The empirical results were examined over different score scales and statistical transformations. The correlation coefficient ranged between 0.61 and 0.78, suggesting that around 36% to 60% of mathematics or science performance can be accounted for by the relationship between these two subjects. (Contains 1 table and 40 references.) (Author/SLD)

Running head: Mathematics/Science Relation

An Analysis of Relationships Between Mathematics and Science Achievement in TIMSS and TIMSS-R

Jianjun Wang

Department of Advanced Educational Studies

California State University, Bakersfield

9001 Stockdale Highway

Bakersfield, CA 93311

Phone: (661) 664-3048

Fax: (661) 664-2016

E-Mail: jwang@csub.edu

Paper presented at the 2003 annual meeting of the American Educational Research Association (AERA), Chicago, IL.

An Analysis of Relationships Between Mathematics and Science Achievement in TIMSS and TIMSS-R

Abstract

Articulation of mathematics and science education is advocated in official documents of several professional organizations. To assess benefit of the curriculum integration, a national indicator needs to be developed from a correlation study of performance scores between the two subjects. In this study, the correlation analysis is conducted at the 8th grade level using international databases from the Third International Mathematics and Science Study (TIMSS) and its repetition in 1999 (TIMSS-R). The empirical results are examined over different score scales and statistical transformations. The correlation coefficient ranges between .61 and .78, and thus, this study seems to conclude that around 36% - 60% of mathematics or science performance can be accounted for by the relationship between these two subjects.

An Analysis of Relationships Between Mathematics and Science Achievement in TIMSS and TIMSS-R

Improvement of student achievement in mathematics and science is part of the *Educate America Act* passed by the U.S. Congress (H.R. 1804). Important documents, such as *Professional Standards for Teaching Mathematics* (National Council of Teachers of Mathematics [NCTM], 1991), *Curriculum and Evaluation Standards for School Mathematics* (NCTM, 1989), *National Science Education Standards* (National Research Council [NRC], 1996), and *Benchmarks for Science Literacy* (American Association for the Advancement of Science [AAAS], 1993) have been developed by professional organizations to strengthen articulation of mathematics and science education. All these national initiatives were built on an assertion that these two subjects were interrelated, and an integrated curriculum may help improve student performance in either subject (see Czerniak, Weber, Sandmann, & Ahern, 1999; Hurley, 2001; Lonning & DeFranco, 1997).

Guided by these national standards, many educators have been involved in curriculum reforms across the nation. "Although there have been numerous curriculum development projects aimed at the integration of science and mathematics education, there has been very little research to evaluate their effectiveness" (Berlin, 1989, p. 74). Ten years later, Miller and Davison (1999) still raised the same question, "What improvements in student learning should be expected from the application of an integrated curriculum?" (p. 29).

In part, the void in this area was caused by lack of assessment data to correlate student performances between mathematics and science. For more than three decades, the National Assessment of Educational Progress (NAEP) has been one of the primary projects to assess the condition of U.S. education. Due to nature of the NAEP design, mathematics and science scores were gathered from different student samples (Allen, Carlson, & Zelenak, 1999). Thus, no students took the NAEP science and mathematics tests concurrently, and no inter-disciplinary analysis can be conducted using the NAEP database.

Besides the domestic projects, large-scale comparative data have been released from the Third International Mathematics and Science Study (TIMSS) and a repeat of the TIMSS project (TIMSS-R) in the late 1990s. Widely cited as an international benchmark in education, the TIMSS and TIMSS-R projects incorporated both mathematics and science tests to measure student academic performance (Martin & Mullis, 1996; Mullis, et al., 2000). Accordingly, an analysis of the score correlations can be conducted in this study to assess the relationship between mathematics and science achievements. Statistical indicators can be developed from the analysis of U.S. 8th grade databases from the TIMSS and TIMSS-R projects. To date, TIMSS/TIMSS-R reports have been largely divided along with subject boundaries (e.g., Beaton et al., 1996a, b; Martin et al., 1998, 2001; Mullis, et al., 1998, 2001). A unique feature of this investigation is represented by a concerted effort to articulate the assessments of student achievement between the two core subjects.

Literature Review

Because student test scores are measured on an interval scale, Pearson correlation coefficient is an appropriate choice to assess the linear relationship between mathematics and science achievements (Ott, 1993). The formula of Pearson r can be written as:

$$r = cov(x_1, x_2)/sqrt[var(x_1)*var(x_2)] \quad (1)$$

Therefore, the computing of Pearson r depends the values of the variances [i.e., $var(x_1)$, $var(x_2)$] and covariance [i.e., $cov(x_1, x_2)$].

Estimation of the variance and covariance parameters hinges on the sampling structure. Built on the NAEP experience, the TIMSS/TIMSS-R projects employed stratified/cluster sampling techniques to facilitate the data collection (Martin, Gregory, & Stemler, 2000; Martin & Kelly, 1997). According to Kish (1965), the assumption of simple random sampling tends to underestimate the variability of statistical estimates for stratified samples. The difference can be described by design effect (deff):

deff=(variance from complex sampling)/(variance from simple random sampling)

The American Institute of Research [AIR] (2003) developed and upgraded a special software package entitled "AM" to analyze data from complex samples. AIR noted,

> AM is a statistical software package for analyzing data from complex samples, especially large-scale assessments such as the National Assessment of Educational Progress (NAEP) and the Third International Mathematics and Science Studies (TIMSS). (http://am.air.org, p. 1)

For the TIMSS/TIMSS-R data, it was reported that a total of five plausible scores have been computed for each student in each subject area, and "one set of the imputed plausible scores can be considered as good as another" (Gonzalez & Smith, 1997, ch. 6, p. 3). The interchangeability of plausible scores suggests equivalency of the *design effect* (deff) among the plausible scores. Under the invariant assumption of the design effect, the AM software is employed to compute correlation coefficients among the plausible scores in each subject.

Research Questions

In the TIMSS and TIMSS-R databases, TIMSS scores have been scaled twice in the 1990s. The old scale was built on the single-parameter Rasch model, and has been used in the original TIMSS reports (Beaton et al., 1996a, b). TIMSS-R employed a three-parameter model from the Item Response Theory (IRT) (Martin et al., 2001; Mullis, et al., 2001). Thus, the TIMSS data have been re-scaled by the three-parameter IRT model to examine the trend between TIMSS and TIMSS-R (Martin, Gregory, & Stemler, 2000). For this reason, plausible scores were computed in the three data files: (1) TIMSS original data, (2) TIMSS-R data, and (3) TIMSS re-scaled data (Gonzalez & Smith, 1997; Martin, Gregory, & Stemler, 2000). To triangulate the research findings from these measures at the 8th grade level, correlational analyses have been conducted in this study to address the following questions:

1. What are the correlation coefficients between mathematics and science achievements using different methods of statistical summary?

2. Are there any differences in the TIMSS correlation coefficients between the new and old scales?

3. Are there any differences in results of the correlational analyses between TIMSS and TIMSS-R?

Methods

After obtaining the correlation coefficients between plausible scores of mathematics and science achievements, an average of these coefficients is needed to present the statistical findings. However, Fisher (1921) raised a caution against the use of a simple average for the r coefficients. He wrote:

> In the neighbourhood of +1, the [correlation coefficient distribution] curves become extremely skew, even for large samples, and change their form so rapidly that the ordinary statement of the <<probable error>> is practically valueless. It was accordingly suggested that the variable r was unsuitable for expressing the accuracy of an observed correlation in these regions but that, by a simple transformation, a variable might be obtained the sampling curves of which are practically normal and of constant standard deviation. (p. 1-2)

Corey, Dunlap, and Burke (1998) concurred, "When correlations come from a matrix, there is a consistent advantage associated with using [Fisher's] z'. Across sample size and numbers of correlations averaged, bias in average r(z)' is smaller than bias in average r" (p. 260).

In this study, the average of r is calculated and compared to the corresponding results from Fisher's (1921) z transformation to check the alternative r estimates between mathematics and science achievements (Question 1). In addition, findings from the new and old TIMSS scales are examined to assess stability of the r estimates (Question 2).

The TIMSS and TIMSS-R results are further analyzed to disentangle the trend of the correlation coefficients at the 8th grade between 1995 and 1999 (Question 3).

Results

Perhaps because of the fairly large sample size involved in the correlation computing, no substantial difference has been found in the results from the average Z score and the Fisher's z transformation (see Table 1). On the other hand, significant differences have been found from the correlation coefficients between the old and new scales in TIMSS. This gap seems to justify need of a scale transformation in the TIMSS and TIMSS-R trend analysis (Martin et al., 2001; Mullis, et al., 2001). On the same new scale of mathematics and science scores, the TIMSS and TIMSS-R results seem fairly consistent with a relative fluctuation less than 4% of the r value. Across all these measures, a moderate to strong degree of relationship ($.61<r<.78$) has been found between mathematics and science achievements.

Discussions

The history of integrating mathematics and science instruction can be traced back to at least the beginning of the 20th century (see Isaacs, Wagreich, & Gartzman, 1997; Lehman & McDonald, 1988). Some educators believe that "The integration of science and mathematics education can provide real world experiences and applications which may encourage student involvement and facilitate the understanding of both science and mathematics concepts, skills, and processes" (Berlin, 1989, p. 73).

Despite the persistent encouragement for curriculum articulation by professional organizations, the connection between mathematics and science may vary in specific subject domains. On one hand, much of physics cannot be properly covered without calling on mathematical concepts and skills. However, the mathematical demand is not as strong in biology, and "other sciences such as psychology might not yet be ready for the kind of mathematization that has taken place in physics" (Orton & Roper, 2000, p. 124).

Consequently, science and mathematics educators have to deal with an "unfocused definition of integration" (Czerniak, Weber, Sandmann, & Ahern, 1999, p. 422). Huntley (1998) proposed a mathematics/science continuum on which both ends represent a clear separation of mathematics and science, and the center represents a compete integration. The TIMSS and TIMSS-R results show an average correlation coefficient between .61 and .78 (see Table 1). Converting to a coefficient of determination (r^2), the results seem to suggest that an integration effort might account for 36% - 60% of mathematics or science performance in the United States according to the international measurements at the 8th grade level. Thus, too much or too little emphasis on the curriculum integration does not seem to have the support from the existing database.

More specifically, the TIMSS instrument also includes some items covering applications of mathematics knowledge in scientific inquiry. For instance, an item on proportional reasoning reads:

L14. The table shows the values of x and y, where x is proportional to y.

x	3	6	P
y	7	Q	35

What are the values of P and Q?

A. $P = 14$ and $Q = 31$

B. $P = 10$ and $Q = 14$

C. $P = 10$ and $Q = 31$

D. $P = 14$ and $Q = 15$

E. $P = 15$ and $Q = 14$

This type of data imputation has been employed in deduction of various models in physics and chemistry, such as Charles' law of thermodynamics (Sears, Zemansky, & Young, 1987). Whereas a random guessing over the five choices could have generated a 20% correct response rate, only 24% eighth grade participants responded correctly to this TIMSS question (http://www.timss.org).

In another example, students were asked to use a classical relationship among time, displacement, and velocity (Hagelberg, 1973). Given four options in the following item, the probability of obtaining a correct answer through guessing is 25%. Across all TIMSS participating nations, only 27% 8th graders answered this question correctly. Hambleton (1988) pointed out, "with difficult multiple-choice tests, a researcher might anticipate considerable guessing on the part of examinees. Needed, therefore, would be a model that could handle this situation" (p. 154).

Q16. How long does it take light from the nearest star other than the Sun to reach Earth?

A. Less than 1 second

B. About 1 hour

C. About 1 month

D. About 4 years

The TIMSS old scale was built on the single-parameter Rasch model, and has resulted in different findings than the one from the TIMSS-R three-parameter IRT model. In general, the Rasch model can be considered as a special case of the three-parameter IRT model under assumptions of *equal item discrimination* and *no correct guessing* among low ability examinees (Hambleton, 1988, Hambleton & Swaminathan, 1985). According to Lange (1997), the TIMSS instrument includes 429 multiple-choice, 43 short-response, and 29 extended-response items. The large number of multiple-choice items, along with the low correct-response rate in this merging area between mathematics and science, seems to support the effort of TIMSS researchers to rescale the TIMSS results on a three-parameter IRT model that has taken the guessing effect into consideration. On the new IRT scale, the correlation coefficient between mathematics and science performances is in a range above .74 and below .78, showing a much higher level of consistency between the TIMSS and TIMSS-R findings (see Table 1).

In summary, although various school initiatives have been introduced across the

United States to integrate mathematics and science curricula (e.g., Judson & Sawada, 2000; Stallings & Ottinger, 1994; Woolnough, 2000), few researchers have examined empirical evidence to disentangle the relationship between the two subject scores. TIMSS and TIMSS-R data provided a unique opportunity to investigate the score correlation using different statistical transformations and measurement scales. Measured by the coefficient of determination (r^2), the data analysis appears to indicate that an integration effort can account for 36% - 60% of mathematics or science performance at the 8th grade level. Joint efforts seem to be needed from mathematics and science educators to further improve student performance on these mathematics-science linkage items beyond the level of random guessing.

References

Allen, N. L., Carlson, J. E., & Zelenak, C. A. (1999). *The NAEP 1996 technical report*. Washington, DC: National Center for Education Statistics.

American Institute of Research [AIR] (2003). *AM statistical software*. [Online] Available at http://am.air.org [March 26, 2003].

American Association for the Advancement of Science [AAAS] (1993). *Benchmarks for science literacy*. New York: Oxford University.

Beaton, A., Martin, M., Mullis, I., Gonzalez, E., Smith, T., & Kelly, D. (1996a). *Science achievement in the middle school years*. Chestnut Hill, MA: Boston College.

Beaton, A., Mullis, I., Martin, M., Gonzalez, E., Kelly, D., & Smith, T. (1996b). *Mathematics achievement in the middle school years*. Chestnut Hill, MA: Boston College.

Berlin, D. F. (1989). The integration of science and mathematics education: Exploring the literature. *School Science and Mathematics*, *89* (1), 73-80.

Corey, D. M., Dunlap, W. P., & Burke, M. J. (1998). Observed and expected bias in average correlation with and without using Fisher's z transformation. Journal of General Psychology, 125, 245-261.

Czerniak, C. M., Weber, W. B., Sandmann, A. & Ahern, J. (1999). A literature review of science and mathematics integration. *School Science and Mathematics*, *99* (8), 421-430.

Fisher, R. A. (1921). On the 'probable error' of a coefficient of correlation deduced from a small sample. *Metron*, *1*, 1-32.

Gonzalez, E. J., & Smith, T. A. (1997). *Users guide for the TIMSS international database*. Chestnut Hill, MA: TIMSS International Study Center.

Hagelberg, M. P. (1973). *Physics*. Englewood Cliffs, NJ: Prentice-Hall.

Hambleton, R. K. (1988). Principles and selected applications of item response theory. In R. L. Linn (Ed), *Educational measurement* (3rd ed.). London: Collier Macmillan.

Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston, MA: Kluwer.

Huntley, M. A. (1998). Design and implementation of a framework for defining integrated mathematics and science education. *School Science and Mathematics, 98* (6), 320-327.

Hurley, M. (2001). Reviewing integrated science and mathematics: The search for evidence and definitions from new perspectives. *School Science and Mathematics, 101* (5), 259-268.

Isaacs, A., Wagreich, P. & Gartzman, M. (1997). The quest for integration: School mathematics and science. *American Journal of Education, 106*, 179-206.

Judson, E. & Sawada, D. (2000). Examining the effects of reformed junior high school science class on students' math achievement. *School Science and Mathematics, 100* (8), 419-425.

Kish, L. (1965). *Survey sampling*. New York: John Wiley & Sons.

Lange, J. D. (1997). *Looking through the TIMSS mirror from a teaching angle*. [On line] Available: http://www.enc.org/topics/timss/additional/documents/0,1341,CDS-

000158-cd158,00.shtm (June 11, 2001).

Lehman, J. R., & McDonald, J. L. (1988). Teachers' perceptions of the integration of mathematics and science. *School Science and Mathematics, 88* (8), 642-649.

Lonning, R. A., & DeFranco, T. C. (1997). Integration of science and mathematics: A theoretical model. *School Science and Mathematics, 97* (4), 212-215.

Martin, M. O., Gregory, K. D., & Stemler, S. E. (2000). *TIMSS 1999: Technical report*. Chestnut Hill, MA: TIMSS International Study Center.

Martin, M. O., & Kelly, D. L. (1997). *Technical report volume II: Implementation and analysis*. Chestnut Hill, MA: TIMSS International Study Center.

Martin, M. O. & Mullis, I. V. S. (1996). Quality assurance in data collection. Chestnut Hill, MA: TIMSS International Study Center.

Martin, M. O., Mullis, I. V. S., Beaton, A. E., Gonzalez, E. J., Smith, T. A., & Kelly, D. L. (1998). *Science achievement in Missouri and Oregon in an international context: 1997 TIMSS benchmarking*. Chestnut Hill, MA: TIMSS International Study Center.

Martin, M. O., Mullis, I., Gonzalez, E., O'Connor, Chrostowski, S., Gregory, K., Smith, T., & Garden, R. (2000) *TIMSS 1999: International science report*. . Chestnut Hill, MA: Boston College.

Martin, M. O., Mullis, I., Gonzalez, E., O'Connor, Chrostowski, S., Gregory, K., Smith, T., & Garden, R. (2001). *Science benchmarking report TIMSS 1999 – Eighth grade*. Chestnut Hill, MA: Boston College.

McBride, J. W., & Silverman, F. L. (1991). Integrating elementary/middle school science and mathematics. *School Science and Mathematics, 91* (7), 285-292.

Miller, K. W. & Davison, D. M. (1999). Paradigms and praxis: The role of science and mathematics integration. *Science Educator, 8* (1), 25-29.

Mullis, I. V. S., Martin, M. O., Beaton, A. E., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1998). *Mathematics achievement in Missouri and Oregon in an international context: 1997 TIMSS benchmarking*. Chestnut Hill, MA: TIMSS International Study Center.

Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Gregory, K. D., Garden, R. A., O'Connor, K. M., Chrostowski, S. J., Smith, T. A. (2000). *TIMSS 1999: International mathematics report*. Chestnut Hill, MA: TIMSS International Study Center.

Mullis, I., Martin, M. O., Gonzalez, E., O'Connor, Chrostowski, S., Gregory, K., Garden, R., & Smith, T. (2001). *Mathematics benchmarking report TIMSS 1999 – Eighth grade*. Chestnut Hill, MA: Boston College.

National Council of Teachers of Mathematics [NCTM] (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

National Council of Teachers of Mathematics [NCTM] (1991). *Professional standards for teaching mathematics*. Reston, VA: Author.

National Research Council [NRC] (1996). *National science education standards*. Washington, DC: National Academy Press.

Orton, T., & Roper, T. (2000). Science and mathematics: A relationship in need of counseling? *Studies in Science Education, 35*, 123-154.

Ott, R. (1993). *An introduction to statistical methods and data analysis* (4th ed.).

Belmont, CA: Duxbury.

Sears, F. W., Zemansky, M. W., & Young, H. D. (1987). *University physics* (7th ed.). Reading, MA: Addison-Wesley.

Stallings, L. & Ottinger, K. (1994). Lessons from landfill. *Science Teacher, 61* (9), 20-25.

Woolnough, J. (2000). How do student learn to apply their mathematical knowledge to interpret graphs in physics? *Research in Science Education, 30* (3), 259-267.

Table 1

Correlation coefficients between mathematics and science achievements at the 8th grade

Project	Measures	N	Average r	Fisher's r(z)
TIMSS	Old scale	7087	.61217	.61220
	New scale	7087	.74936	.74937
TIMSS-R	New scale	9072	.77766	.77763

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC
Educational Resources Information Center

REPRODUCTION RELEASE

(Specific Document)

TM034833

I. DOCUMENT IDENTIFICATION:

Title: An analysis of relationship between mathematics and science achievement in TIMSS and TIMSS-R	
Author(s): Jianjun Wang	
Corporate Source: American Educational Research Association, Chicago, IL	Publication Date: April, 2003

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1 ☑	Level 2A ☐	Level 2B ☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Signature: [illegible]	Printed Name/Position/Title: Jianjun Wang, Professor	
Organization/Address: California State University, Bakersfield	Telephone: [illegible]	FAX: [illegible]
	E-Mail Address: [illegible]	Date: [illegible]