An Empirical Study of Relationships

Between Student Self-Concept and Science Achievement in Hong Kong

Jianjun Wang
California State University, Bakersfield

Steve Oliver
University of Georgia

Augustine Garcia
California State University, Bakersfield

Paper presented at the 2004 annual meeting of the American Educational Research Association (AERA), San Diego, CA.

An Empirical Study of Relationships

Between Student Self-Concept and Science Achievement in Hong Kong

Abstract

Positive self-concept and good understanding of science are important indicators of scientific literacy endorsed by professional organizations. The existing research literature suggests that these two indicators are reciprocally related and mutually reinforcing. Generalization of the reciprocal model demands empirical studies in different cultures. In this investigation, Hong Kong has been identified as a meeting place between the West and the East, and TIMSS and TIMSS-R databases are analyzed to confirm features of the reciprocal model two years *before* and *after* Hong Kong's political transition in 1997. The result may not only help generalize the reciprocal model in a different culture setting, but also assess contextual changes in the language of instruction from English to Chinese brought by the political transition.

An Empirical Study of Relationships

Between Student Self-Concept and Science Achievement in Hong Kong

Self-concept is an intrinsic characteristic that has a profound impact on student future success (Goleman, 1995; McClelland, 1993; Sternberg & Wagner, 1993). In particular, a person's belief in his or her ability to do well in science is recognized as a key component of scientific literacy by science educators (AAAS, 1990; Wilkins, 2003). Marsh, Hau, and Kong (2002) further noted,

> In academic self-concept research, support for the main theoretical models has been based largely on responses by students from Western countries, particularly English-speaking students in Australia, Canada, and the United States. (p. 728)

Accordingly, it is important to examine the relationship between self-concept and academic achievement in various culture settings.

Hong Kong is a unique place harboring different cultures between the West and the East. While the Chinese majority speaks the Cantonese (a local Chinese dialect) as a mother tongue, English is valued highly as a language medium of international trade and academic discourse. During much of the post-war period in Hong Kong, the vast majority of students received a primary education in Chinese-medium schools, and then switched to Anglo-Chinese schools for their secondary education. Hau, Kong, Marsh, and Cheng (2000) observed that "Hong Kong is an ideal setting for testing the juxtaposition of self-concept in native and nonnative languages as both Chinese (the native language) and English (the non-native language) are considered extremely important in the high school curriculum and society" (p. 5).

On July 1, 1997, Hong Kong terminated its 150-year British colonial history, and reunited with China. The new local government designated a transition period to promote

instruction in Chinese/Cantonese in most secondary schools (Evans, 2000). At this historical conjecture, Hong Kong participated in the Third International Mathematics and Science Study (TIMSS) in 1995 and a repeat of TIMSS (named TIMSS-R) in 1999. The purpose of this investigation is to study the relationship between student self-concept and science achievement two years *before* and *after* the sovereignty switch using the TIMSS and TIMSS-R databases. According to Marsh et al. (2002), "Previous research suggests that Chinese students differ from Western students in ways that may be relevant to how they construct their self-concepts" (p. 728). Therefore, this empirical data analysis may not only help disentangle the relationship between self-concept and science achievement in different culture settings, but also assess changes in the language of instruction from English to Chinese brought by the political transition.

Literature Review

Development of individual self-concept can be dated back to Socrates' call of "knowing thyself" more than 2000 years ago (see Hamachek, 2000). In the last century, behavior scientists attempted to conceptualize the idea of self in multiple ways (Cooley, 1902; James, 1950; Mead, 1934). In an attempt to connect self-concept with academic achievement, Shavelson, Hubner, and Stanton (1976) proposed a multifaceted, hierarchical model from an extensive literature review. Apex of the hierarchical structure was split into academic and nonacademic components of self-concept. The relationship between academic self-concept and academic achievement has been postulated to be reciprocal and mutually reinforcing (Marsh et al., 1999). Marsh et al. (2002) reported that "the results of previous research provide general support for a reciprocal effects model" (p. 729).

Nonetheless, more recent cross-cultural investigations have revealed inconsistent findings from international databases (see Wilkins, 2003). For instance, Wilkins, Zembylas, and Travers (2002) analyzed the TIMSS international data, and reported a positive relationship between self-concept and science achievement for 16 different countries. Meanwhile, Kifer's (2002) analysis suggested that many of the highest performing countries had some of the lowest overall beliefs in student self-ability. Incorporating considerations of the instructional language, Evans (2002) further noted needs to articulate the relationship between self-concept and student achievement in specific subject areas.

Nonetheless, "The role of language of instruction in the formation of a positive academic self-concept has received surprisingly little attention in the research into immersion, bilingualism, and second-language instruction" (Marsh, Hau, & Kong, 2002, p. 733). On one hand, researchers tend to believe that students learn more effectively when taught through their mother tongue (see Cummins, 1996; Garcia, 1993). The improvement of achievement may in turn enhance positive attitude, as suggested by the reciprocal model (Marsh et al., 1999). On the other hand, schools that support instruction in English are likely to be the most prestigious and academic ones in Hong Kong. Thus, students may have established positive self-concepts in this setting as well. The western literature in this area seemed to suggest potential "cultural ambivalence" from replacement of the first language by the second language (e.g., Krashen, 1997; Ogbu, 1992, 1999), but that concern was largely hinged on consideration of cultural and language protections for minority groups. In contrast, the immersion in English instruction was a preferred choice of the majority parents in Hong Kong, and has been positively valued by the society. Therefore, the western theories on the negative impact might not fit the specific situation in Hong Kong.

To date, besides a couple of articles developed by Marsh and his colleagues (i.e., Hau, Kong, Marsh, & Cheng, 2000; Marsh, Hau, & Kong, 2000 & 2002), few other researchers examined the relationship between self-concept and science achievement in Hong Kong's context. The data analyzed by Marsh and his colleagues were gathered from 56 high schools under guidance of Hong Kong's Educational Commission in 1995, an era before of the sovereignty switch (see Marsh, Hau, & Kong, 2000, p. 311). To assess impact of the political transition, more recent data need to be analyzed to reflect changes in Hong Kong before and after July 1, 1997.

In summary, this study is designed to disentangle the relationships between student self-concept and science achievement in an era of adjusting English status as a medium of instruction in Hong Kong. Under a "one-country, two-system" framework, Hong Kong remains to be a democratic society and the perceived English importance by education stakeholders plays a crucial role in choosing an instructional language in school. Empirical data on English importance has been gathered in TIMSS and TIMSS-R surveys before and after the political transition. Because most theories of bilingual education were developed in western cultures for minority groups, this investigation not only addresses a critical education topic for the majority residents in Hong Kong, but also facilitates enrichment of the existing literature in bilingual and science education.

Research Questions

Marsh et al. (1999) reviewed the existing literature, and recommended multilevel analyses to examine generalization of the reciprocal modeling across various schools. Research questions that guide this investigation are:

1. What proportion of the variance in science achievement has been distributed at the student and school levels?
2. What is a plausible model of student self-concept and science achievement under the context of reforming English role in instruction during the political transition?
3. What are differences and/or similarities of the empirical model reconfirmed by the TIMSS and TIMSS-R data analyses?

Methods

Variable Selection

Science scores have been gathered from the 8th grade level in TIMSS and TIMSS-R projects. To avoid a low response rate, each student was tested on a subset of science items, and a total of five plausible scores have been imputed to represent the overall student achievement. According to Gonzalez and Smith (1997), "one set of the imputed plausible scores can be considered as good as another" (ch. 6, p. 3). To enhance comparability between the TIMSS and TIMSS-R results, TIMSS scores have been rescaled at the 8th grade level using the TIMSS-R procedure, and the data are released at a public website (http://www.timss.org). Plausible scores from TIMSS and TIMSS-R are employed in this study to indicate student achievement in the 8th grade science.

Selection of self-concept indicators is also grounded on the research literature. Psychologists suggested that self-concept has two major aspects, "the self as a doer" and "the self-as-object" (Hamachek, 2000; James, 1890). From the doer's perspective, TIMSS and TIMSS-R gathered "Students' self-perceptions about usually doing well in science" (Supplement 2 of the TIMSS User Guide, section 1, p. 8). Regarding the self as an object, students had a

chance to express their feeling of getting bored by science. In addition, students reported importance of learning English as perceived by self, friends, and parents, and thus, the push for English acquisition can be examined before and after the political transition. To facilitate interpretation of the research findings, these indicators have been scaled in such a way that a higher value represents a more positive response in each dimension.

Statistical Modeling

While variables of self-concept and science achievement were measured at the student level, their relationship could have been influenced by school characteristics. It is prudent to partition the score variance at student and school levels (Question 1), and assess the feasibility of generalizing research findings across different schools. In this study, the SAS PROC MIXED routine has been employed to complete the variance partition for each of the five plausible scores from TIMSS and TIMSS-R.

The medium of instruction is an important issue related to science education. In Hong Kong, Tao (1994) projected that "The problem is particularly acute in science since the language used has a characteristic formal style" (p. 332). Accordingly, the reciprocal model should incorporate student reported language push from self, friends, and parents (Question 2), and multiple indicators have been employed to identify each latent factor (Figure 1). "Today, it is commonly accepted that multiple observed variables are preferred over a single variable in defining a latent variable" (Schumacker & Lomax, 1996, p. 55). The use of multiple indicators permits an assessment of measurement errors in this study, and goodness-of-fit indices have been computed from the TIMSS and TIMSS-R databases for reconfirmation of the reciprocal model (Question 3).

To keep Figure 1 more readable, the measurement errors considered in this investigation are not depicted along with factor loadings, path coefficients, and correlation coefficients.

_____

Insert Figure 1 around here

_____

Results

The multilevel analysis revealed that score variances were much smaller at the school level than at the student level (Table 1).

_____

Insert Table 1 around here

_____

Goodness of fit indices are listed in Table 2 to assess plausibility of fitting the reciprocal model (Figure 1) to the TIMSS and TIMSS-R databases.

_____

Insert Table 2 around here

_____

Researchers noted that results of statistical testing are sensitive to sample sizes, and trivial differences might have a significant level simply because of a large sample size (see Henson & Smith, 2000). To avoid this problem in statistical reporting, effect sizes as represented by the real value differences have been recommended by the American Psychological Association for dissemination of statistical findings (see Thompson, 1998). Table 3 contains parameter estimates for the reciprocal model in Figure 1 using the TIMSS and TIMSS-R databases. The effect sizes of the structural relationships before and after Hong Kong's political transition are presented in Table 4 to compare the model difference.

Insert Tables 3 & 4 around here

Discussion

As was mentioned earlier, TIMSS and TIMSS-R were gathered four years apart, and in between there is a historical handover of Hong Kong from Britain to China in 1997. Because similar sampling strategies and data collection techniques were employed in these projects, potential impact of the historical transition can be assessed through analyses of the 1995 and 1999 databases.

A political condition for the sovereignty switch is to allow Hong Kong to retain its existing political system. The China government did not impose specific policy changes on the Hong Kong residents. The effect sizes in Table 4 reflect changes in the structure of the reciprocal modal during the transition period. The small effect sizes seemed to confirm stability of the proposed model on the time dimension (Figure 1).

In addition, Hong Kong has an urban environment with Chinese accounting for 98% of the population, and a strong commitment to child education, as part of Chinese tradition, is commonly shared in almost all families. Because of the lack of ethnic variation in the school population, it is generally understandable that score variances were much smaller at the school level than at the student level (Table 1).

Nevertheless, Table 1 also showed an increase in the school-level variability from TIMSS to TIMSS-R. In part, the larger score variability at the school level could be resulted from a new government policy implemented in 1997 that separated school instruction on two tracks by Chinese or English language of instruction. Recent studies of language use in English

schools revealed that the use of Cantonese (i.e., an oral Chinese language) has become widespread among Hong Kong English teachers for various reasons (Bray, 1997; Lin, 1990; Pennington, 1995, 1999). The mixed code instruction has somehow blurred the difference between English- and Chinese-medium schools. The language labeling might have limited teacher application of mixed codes between Chinese and English originally practiced in both tracks, and sharpened the difference at the school levels after the political transition.

This speculation is also supported by results of factor loadings from the measurement of science achievement. The plausible scores have shown similar factor loadings in TIMSS or TIMSS-R (see $\lambda_{y1} \ldots \lambda_{y5}$ in Table 3), which concurs equivalency of the plausible scores in representing student achievement in science (Gonzalez & Smith, 1997). On the other hand, higher factor loadings have been obtained from TIMSS than from TIMSS-R. Factor loadings indicate a simple correlation between the plausible scores and the underline factor of science achievement (Figure 1). Before implementing the language tracking system in 1997, more coherent

Reconfirm the results by factor loadings (besides equivalency). lower loading

Self is more than peer and parent – source of response

Almost no reciprocal relation

More impact on achievement

Increase doer's loading

Among a variety of model fitting indices, Joreskog and Sorbom (1981) advocated the use of goodness-of-fit index (GFI) to measure the relative amount of variances and covariances commonly explained by the model. Marsh, Balla, and McDonald (1988) suggested the adoption of root mean square residual (RMR) to "justify the conclusion that a model adequately fits a

particular set of data" (p. 391). Table 2 contains GFI and RMR indices from fitting the reciprocal model with or without gender separation. The use of multiple fitting indices follows Bollen's (1989) recommendation to confirm the model-data-fit through result triangulations. The small RMR and large GFI values consistently suggest a good fit of the reciprocal model to the TIMSS and TIMSS-R databases.

because "language is an important cultural factor when comparatively assessing students who speak, read, write, and listen using entirely different communication system" (Holliday & Holliday, 2003, p. 252),

4. What proportion of the variance in science achievement has been distributed at the student and school levels?
5. What is a plausible model of student self-concept and science achievement under the context of reforming English role in instruction during the political transition?
6. What are differences and/or similarities of the empirical model reconfirmed by the TIMSS and TIMSS-R data analyses?

Because English is a language largely confined within professional domains, Hong Kong residents rarely use it in their daily communications (Tao, 1994). Consequently, labeling schools according to their language use does not completely reflect the process of instruction. Given the dependency of science education on language acquisition and mathematical preparation, empirical data from TIMSS and TIMSS-R need to be analyzed to disentangle the impact of English emphasis as perceived by education stakeholders, such as students, peers, and parents.

. Factor loadings in Table 3 indicate contributions of each indicator variables on the corresponding latent factors in Figure 1. The path coefficients suggest that importance of English acquisition seems to have stronger impact on science achievement and less impact on science self-concept. However, the reciprocal relation between the achievement and self-concept appears to be much weaker (Table 3). Due to implementation of a "one China, two system" policy, the effect sizes for the period of 1995-1999 political transition are fairly small (Table 4). The model fitting indices in Table 2 show a strong support for the theoretical model depicted in Figure 1.

As a result, the new Hong Kong government had to modify its original policy on the language labeling, and let Chinese-medium schools conduct classes in English (Education lacks a pass mark despite reforms, Hong Kong Imail, October 6, 2000, p. 1). This change was primarily due to parental push for retaining English instruction during the period of Hong Kong's political transition.

References

American Association for Advancement of Science (1990). Science for all Americans. New York: Oxford University Press.

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

Bollen, K. A. (1989). Structural equations with latent variables. New York: Wiley.

Bray, M. (1997). Education and colonial transition: The Hong Kong experience in comparative perspective. Comparative Education, 33 (2), 157-169.

Cooley, C. H. (1902). Human nature and the social order. New York: Scribner.

Evans, S. (2000). Hong Kong's new English language policy in education. World Englishes, 19 (2), 185-204.

Evans, S. (2002). The medium of instruction in Hong Kong: Policy and practice in the new English and Chinese streams. Research Papers in Education, 17 (1), 97-120.

Goleman, D. (1995). Emotional intelligence. New York: Bantam.

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

Hamachek, D. (2000). Dynamics of self-understanding and self-knowledge: Acquisition, advantages, and relation to emotional intelligence. Journal of Humanistic Counseling, Education and Development, 38 (4), 230-242.

Hau, K. T., Kong, C. K., Marsh, H. W., & Cheng, Z. J. (2000, April). Extension of the internal/external frame of reference model of self-concept formation: Importance of native and nonnative languages for Chinese students. Paper presented at the annual meeting of American Educational Research Association, New Orleans, LA.

Henson, R. K. & Smith, A. D. (2000). State of the art in statistical significance and effect size reporting: A review of the APA task force report and current trends. Journal of Research and Development in Education, 33 (4), 285-296.

James, W. (1890). Principles of psychology (Vol. 1). New York: Holt.

Joreskog, K. G., & Sorbom, D. (1981). Analysis of linear structural relations by the method of maximum likelihood. Chicago, IL: International Educational Services.

Kifer, E. A. (2002). Students' attitudes and perceptions. In D. F. Robataille & A. E. Beaton (Eds.), Secondary Analysis of the TIMSS Results: A Synthesis of Current Research. Boston, MA: Kluwer Academic Publishers.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. Psychological Bulletin, 103, 391-410.

Marsh, H. W., Byrne, B. M., & Yeung, A. S. (1999). Causal ordering of academic self-concept and achievement: Reanalysis of a pioneering study and revised recommendations. Educational Psycholigist, 34 (3) 155-167.

Marsh, H. W., Hau, K. T., & Kong, C. K. (2002). Multilevel causal ording of academic self-concept and achievement: Influence of language of instruction (English compared with Chinese) for Hong Kong students. American Educational Research Journal, 39 (3), 727-763.

McClelland, D. C. (1993). Intelligence is not the best predictor of job performance. Current Directions in Psychological Science, 2, 5-6.

Mead, G. H. (1934). Mind, self, and society. Chicago: University of Chicago Press.

Schumacker, R. E. & Lomax, R. G. (1996). A beginner's guide to structural equation modeling.

Mahwah, NJ: Lawrence Erlbaum.

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-concept: Validation of construct interpretations. Review of Educational Research, 46, 407-441.

Sternberg, R. J., & Wagner, R. K. (1993). The g-eocentric view of intelligence and job performance is wrong. Current Directions in Psychological Science, 2, 1-5.

Tao, P. K. (1994). Comprehension of non-technical words in science: The case of students using a 'foreign' language as the medium of instruction. Research in Science Education, 24, 322-330.

Thompson, B. (1998). Statistical significance and effect size reporting: Portrait of a possible future. Research in the Schools, 5 (2), 33-38

Wilkins, J. L. M. (2003, April). The relationship between mathematics and science self-concept and related achievement: An international investigation. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Wilkins, J. L. M., Zembylas, M., & Travers, K. J. (2002). Investigating correlates of mathematics and science literacy in the final year of secondary school. In D. F. Robataille & A. E. Beaton (Eds.), Secondary Analysis of the TIMSS Results: A Synthesis of Current Research. Boston, MA: Kluwer Academic Publishers.

Table 1

Ratio of variances in science plausible scores (PS) distributed at school and student levels[1]

Project	PS1	PS2	PS3	PS4	PS5	Mean
TIMSS	.027	.029	.028	.029	.028	.028
TIMSS-R	.049	.044	.049	.047	.040	.045

[1] The ratio is computed by dividing the partitioned score variance at the school level over the corresponding variances at the student level.

Table 2

Model fitting indices for the TIMSS and TIMSS-R databases

	TIMSS Sample	TIMSS-R Sample
RMR	0.07	0.07
GFI	0.99	0.98

Table 3

Parameter estimates for the reciprocal model using the TIMSS and TIMSS-R databases

Estimates	TIMSS	TIMSS-R
Factor Loadings		
λ_{y1}	0.93	0.85
λ_{y2}	0.93	0.87
λ_{y3}	0.93	0.87
λ_{y4}	0.93	0.87
λ_{y5}	0.93	0.87
λ_{y6}	0.47	0.61
λ_{y7}	0.59	0.59
λ_{x1}	0.65	0.61
λ_{x2}	0.61	0.63
λ_{x3}	0.76	0.84
Reciprocal Relation		
β	0.07	0.07
Path Coefficients		
γ_1	0.29	0.28
γ_2	0.25	0.24

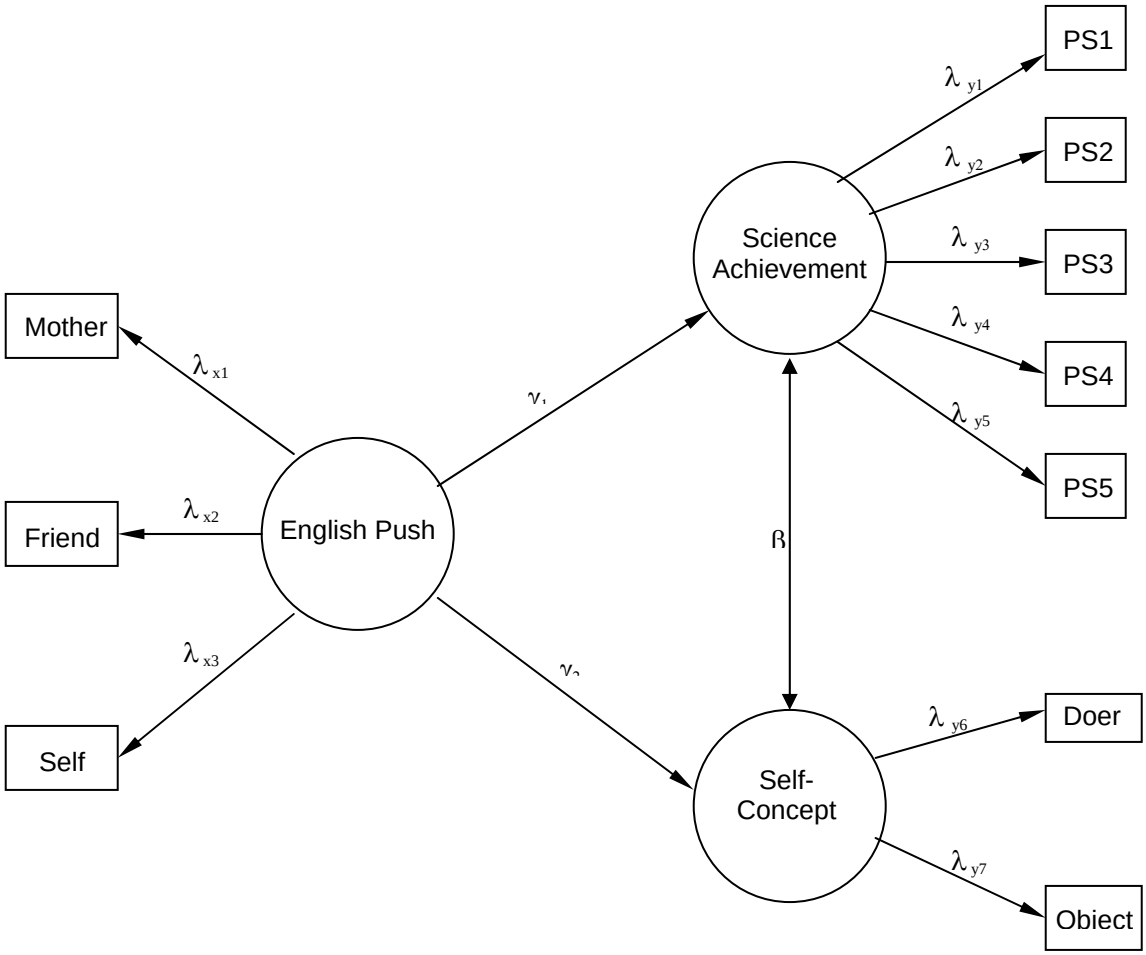
Table 4

Effect sizes of the structural relations from the period of Hong Kong's political transition

Structural Parameter	β	γ_1	γ_2
Effect Size	0	0.01	0.01

Figure 1

A Structural Model of Self-Concept and Science Achievement from TIMSS



Notes:

1. PS1, … PS5 are plausible science scores imputed from TIMSS or TIMSS-R projects under a three-parameter item response theory (IRT) model.
2. The “doer” and “object” indicators are based on student responses on whether they can do well in science and felt bored about science.
3. Indicators for the “English push” factor are based on student responses on importance of learning English perceived by self, friends, and parents.