Extension of the Internal/External Frame of Reference Model of Self-Concept Formation:

Importance of Native and Nonnative Languages For Chinese Students

Kit-Tai Hau, The Chinese University of Hong Kong

Chit-Kwong Kong, Yenching College

Herbert W. Marsh, University of Western Sydney

24 February, 2000

Running Head: **Internal/External Frame of Reference Model**

Author Note

Paper presented at the American Educational Research Association Annual Meeting, April 2-28, New Orleans. This longitudinal study was part of a much larger educational policy research project evaluating secondary schooling in Hong Kong. The academic achievement data has been graciously provided by the Educational Research Section of the Hong Kong Education Department to the authors for this research. The research was funded in part by grants from the Australian Research Council and the Hong Kong Institute of Educational Research. Requests for further information about this investigation should be sent to Kit-Tai Hau, Faculty of Education, The Chinese Univeristy of Hong Kong, Shatin, N.T., Hong Kong; E-mail: KTHAU@CUHK.EDU.HK.

Abstract

We extend the internal/external frame of reference (I/E) model of self-concept formation by relating Chinese, English, and math achievement to Chinese, English, and math self-concepts in a five-year longitudinal study based on a large (N = 9,482) representative sample of Hong Kong high school students. Tests of the I/E model are typically based on math and English constructs for a single wave of data in Western countries. We extend this research, testing its cross-cultural generalizability to a non-Western country, including native and nonnative languages as well as mathematics, and evaluating longitudinal effects over a five-year period starting shortly before the beginning of high school. In support of the extended I/E model: a) math, English, and Chinese achievements were highly correlated whereas math, English, and Chinese self-concept were nearly uncorrelated; b) math, English, and Chinese achievements each had positive effects on the matching self-concept domain, but negative effects on nonmatching domains (e.g., English achievement had a positive effect on English self-concept but negative effects on math and Chinese self-concepts); and c) these results were very stable over time.

Cognition, learning, achievement, academic success, and the self are central concepts in educational psychology. The self, academic self-concepts that students hold of themselves, and achievement-related cognitions are playing an increasingly important role in cognitive models of the student as an active learner. In this study we evaluate and extend a process model of self-concept formation that posits a, perhaps, unexpected pattern of relations between academic achievement in different school subjects and the corresponding academic self-concepts in these specific school subjects. In this model, academic self-concept in specific academic domains is much more than a somewhat noisy indicator of academic achievement. Students, in forming their academic self-concepts, must evaluate their academic accomplishments in relation to some standards or frames of reference. Even if students achieve similar accomplishments, their academic self-concepts will differ if they have different frames-of-reference. Pertinent self-concept assessment and appropriate classroom practice depends upon teachers being able to interpret these differing self-perception processes accurately. Theoretical models posit the cognitive basis that students use to determine appropriate frames of reference in the formation of their academic self-concepts. Because such frame-of-reference effects might be specific to particular cultural settings, it is useful to test the generalizability of predictions based in one culture in different cultures. Here we describe the theoretical basis and empirical support for the internal/external frame of reference (I/E) model that is based primarily on studies in Western countries. In this model, it is proposed that students compare their own accomplishments in different academic subjects (i.e., I am good at math because my math accomplishments are better than my English accomplishments) in addition to the more typical process of comparing their own accomplishments with those of other students or other external standards. In this study, we test the generalizability of this support for the I/E model in a Hong Kong sample of high school students, extend the model by considering self-concepts and academic achievement in relation to both Chinese (the native language) and English (a non-native language that is highly valued) as well as mathematics, and extend the model by testing the strength of effects over five years using longitudinal data starting shortly prior to entry into high school.

Internal/External Frame Of Reference (I/E) Model

The I/E model (for further discussion, see Marsh, 1986, 1990a, 1993; Marsh, Byrne, & Shavelson, 1988; Skaalvik & Rankin, 1995) was initially developed to explain why math and verbal self-concepts are almost uncorrelated even though corresponding areas of academic achievement are substantially correlated (typically .5 to .8, depending on how achievement is measured). Individuals who are good in one area tend to be good in the other. Verbal and Math self-concepts, however, are nearly uncorrelated. People think of themselves as "math" persons or "verbal" persons. According to the I/E model, academic self-concept in a particular school subject is formed in relation to two comparison processes or frames of reference. The first is the typical external (normative) reference in which students compare their self-perceived performances in a particular school subject with the perceived performances of other students in the same school subject and other external standards of actual achievement levels. If they perceive themselves to be able in relation to other students and objective indicators of achievement, then they should have a high academic self-concept in that school subject. The second is an internal (ipsative-like) reference in which students compare their own performance in one particular school subject with their own performances in other school subjects. If, for example, mathematics is their best school subject, then they should have a positive math self-concept that is higher than their verbal self-concept.

To clarify how these two processes operate, consider a student who accurately perceives him or herself to be below average in both verbal and math skills (an external comparison), but who is better at mathematics than verbal and other school subjects (an internal comparison). The student's math skills are below average relative to other students and objective indicators of math achievement (the external comparison), and this should lead to a below average math self-concept. However, this student's math skills are above average relative to his or her other school subjects (an internal comparison) and this should lead to an above-average math self-concept. Depending upon how these two processes are weighted in the formation of self-concept, this student may have an average or even above-average math self-concept even though he or she has below-average math skills. Similarly, a student who is very bright in all school subjects may have an average or even below-average math self-concept if the student perceived mathematics to be his or her worst subject.

The external comparison process should result in substantial positive correlations between math and verbal self-concepts because math and verbal achievements are substantially positively correlated. The ipsative, internal comparison process should result in a negative correlation between math and verbal self-concepts because the average correlation among ipsative scores is necessarily negative (i.e., an increase in any one score must result in the counterbalancing decrease in average of the remaining scores if they are ipsative). Self-concept responses, however, are affected by both these processes. Hence, the joint operation of these processes, depending on the relative weight given to internal and external comparisons, is consistent with the near-zero correlation between math and verbal self-concepts that led to the development of the I/E model. It is, however, important to emphasize that support for the I/E model does not require the correlation between math and verbal self-concepts to be zero, but only that it be substantially less than the typically substantial correlation between math and verbal achievement.

Insert Figure 1 About Here

Stronger tests of the I/E model are possible when math and verbal achievements are related to math and verbal self-concepts (see Figure 1A). The external comparison process predicts that good math skills lead to higher math self-concept and that good verbal skills lead to higher verbal self-concept. According to the internal comparison process, however, good math skills should lead to lower verbal self-concept (once the positive effects of good verbal skills are controlled). The better I am at mathematics, the poorer I am at verbal subjects (relative to my good math skills). Similarly, better verbal skills should lead to lower math self-concept (once the positive effects of good math skills are controlled). In models used to test this prediction (Figure 1A), the paths leading from math achievement to math self-concept and from verbal achievement to verbal self-concept (the gray horizontal lines in Figure 1A) are predicted to be substantially positive (indicated by "++" in Figure 1A). However, the paths leading from math achievement to verbal self-concept and from verbal achievement to math self-concept (the dark lines in Figure 1A) are predicted to be negative (indicated by "– " in Figure 1A). Support for these predictions comes from a large body of research based on a variety of different self-concept measures and from a variety of different countries (Marsh, 1990a; Marsh, Byrne, & Shavelson, 1988; Marsh & Craven, 1997; Marsh & Yeung, 1998; Skaalvik & Rankin, 1995).

Extensions of the Internal/External Frame of Reference Model

A number of researchers have proposed various extensions of the I/E model (e.g., Bong, 1998; Marsh & Köller, 1999; Skaalvik & Rankin, 1995; Yeung & Lee, 1999). In particular, a number of studies conducted in different countries where Verbal self-concept is in relation to a native language other than English (Norwegian, Skaalvik & Rankin, 1995; German, Marsh & Köller, 1999; Chinese, Yeung & Lee, 1999). In a study particularly relevant to the present investigation, Bong (1998) tested the I/E model using a broader range of academic domains, including measures of self-concept and achievement in six school subjects. In her original analysis, the verbal constructs of self-concept and achievement were based on English, Spanish, and American history, whereas the mathematics constructs of self-concept and achievement were based on algebra, geometry, and chemistry. Bong found only partial support for the I/E model. In a reanalysis of this data, Marsh, Bong, and Yeung (1999) reported that Spanish self-concept was nearly uncorrelated with other verbal self-concept factors, and that the corresponding Spanish achievement was only moderately related to either verbal or mathematics achievements. Noting that a majority of the students in the study were native Spanish speakers living in Metropolitan Los Angeles, the authors suggested that the students might view English and Spanish self-concepts as very distinct. Following from this reasoning, they proposed a post hoc model with Spanish, English, and mathematics self-concepts and achievements instead of the typical I/E model based on only English and Math constructs. Extending the logic of the I/E model, they demonstrated that (a) Spanish achievement had a positive effect on Spanish self-concept, but negative effects on mathematics and English self-concepts, (b) mathematics had a positive effect on mathematics self-concept, but negative effects on English and Spanish self-concepts, and (c) English achievement had a positive effect on English self-concept, but negative effects on Math and Spanish self-concepts. Marsh et al. (1999) concluded that the results provided strong support for the original I/E model based on math and verbal constructs, and a potentially important extension of the model when native and nonnative languages were both included in the model. An important implication of the Marsh et al. (1999) study is the juxtaposition of the self-concepts of native and nonnative languages. Their results suggested that self-concepts in native and nonnative languages are likely to be very distinct from each other and from math self-concept, and to have distinct patterns of relations with corresponding measures of achievement.

The Present Investigation

In our five year longitudinal study we evaluate the extensions of the I/E model for a large (N = 9,482) representative sample of Hong Kong high school students. This allows us to evaluate the generalizability of predictions based on the I/E model in a non-Western culture. The Chinese culture is low on the cultural value of individualism and high on collectivism (e.g., Bond, 1996; Diener, Diener, & Diener, 1995; Hofstede, 1991). Thus, Hong Kong students may be less susceptible to the social comparison processes than those in individualistic settings (e.g., McFarland & Buehler, 1995; also see

Markus & Kitayama, 1991; Triandis, 1989). Furthermore, consistent with this potential deemphasis of social comparison processes, Hau and Salili (1991, 1996) found that Hong Kong students attribute their examination results more to effort than to ability and that they concentrate more on their own improvement over time than on comparisons with other students as determinants of perceived academic achievement.

The I/E model posits an important pattern of relations among variables that are typically collected on a single occasion. This reliance on a single wave of data, however, undermines our ability to evaluate the temporal aspects of this process. In the present investigation, the achievement measures were standardized test scores collected prior to the start of high school (T0, in 6^{th} grade). These scores (provided by the Hong Kong Department of Education) are the basis on which elementary students are allocated to different Hong Kong high schools and, thus, are very important to students, their parents, and their schools. Hence, they are likely to be very influential in terms of the formation of academic self-concept and to have potentially long-lasting effects. The three waves of self-concept measures were collected approximately two, three, and four years after the collection of the achievement measures (in 8^{th}, 9^{th}, and 10^{th} grades). Hence, this allows us to chart the strength of support for predictions based on the I/E model over this critical period during the early years of high school. In particular, we evaluate the extent to which the effects of the achievement measures collected prior to the start of high school become weaker over time in separate analyses of each wave of self-concept data. Then, by combining all three waves of self-concept data into a single analysis, we are able to test the extent to which support for the I/E effects at one time are mediated by I/E effects on earlier measures of academic self-concept.

Most importantly, this Hong Kong data provides an ideal basis to test speculations by Marsh et al. (1999) about the extension of the traditional I/E model (see Figure 1B). More specifically they proposed that native and non-native languages form separate areas of self-concept that provide distinct bases for forming self-concepts in other academic areas. 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 as extremely important in the high school curriculum and society. Hong Kong was a British colony but became a special administrative region of the People's Republic of China in July 1, 1997. The majority (98%) of the people in Hong Kong are Chinese and most of them use Chinese (Cantonese) as their mother tongue and in their daily speech. On the other hand, as a former British colony and an international center of business, English is used widely in the government and the business sectors. Thus, a good mastery of English provides enhanced opportunities for higher education and better jobs and, hence, upward social mobility. Hence, English language is strongly valued in the Hong Kong school curriculum. In summary, there are strong practical implications in examining the juxtaposition of self-concepts in English (the non-native language) and in Chinese (the native language) for Hong Kong secondary students. Following from the Marsh et al. (1999) extension of the I/E model (see Figure 1B), we predict that:

1. Correlations among the Chinese, English, and Math self-concepts are very small (close to zero) and substantially less than the large correlations among corresponding Chinese, English, and Math achievement scores;
2. Mathematics achievement has a large positive effect on Math self-concept, but smaller negative effects on English and Chinese self-concepts;
3. English achievement has a large positive effect on English self-concept, but smaller negative effects on Math and Chinese self-concepts; and
4. Chinese achievement has a large positive effect on Chinese self-concept, but smaller negative effects on Math and English self-concepts.

Method

Sample

Our study is part of a large-scale investigation of secondary schooling in Hong Kong. The sample consists of all students (N = 9,482) in the 55 high schools who completed the SDQ-II in 8^{th} grade (T2), the first year during which self-concept responses were collected. The sample includes a diversity of schools that is broadly representative of Hong Kong secondary schools in terms of religious background, mode of government subsidy, and gender grouping.

Procedures and Measures

Achievement Scores in Chinese, Mathematics, and English. In Hong Kong, all 6^{th} grade students are allocated a moderated placement score that represents an internal aggregate of achievement in all school subjects except physical education (although Chinese, English and mathematics are weighted more heavily) that is moderated by external examinations. The external examinations are standardized measures of general ability with separate mathematics and verbal (Chinese) components that are administered by the Department of Education. These "high stakes" scores are the primary basis for the extremely important selection into high schools. For present purposes, achievement scores consist of standardized test scores (based on the moderator examination) for Chinese and mathematics, and a moderated school-based performance measure of English. These scores were provided for all students by the Hong Kong Department of Education for purposes of the present investigation. We refer to 6^{th} grade as T0 because it is prior to the start of high school in 7^{th} grade.

At T1, T2, and T3, students also completed achievement tests administered by the Hong Kong Department of Education according to a modified random matrix sampling design. Each student was randomly assigned an achievement test in only one of the three core subjects (Chinese, English, and mathematics). Thus, for example, only 1/3 of the students completed a math achievement test in any one of the three years, only a small proportion completed a math test in all three years (approximately 1/3 x 1/3 x 1/3 = 4%), and many did not complete a math test in any of the three years (approximately 2/3 x 2/3 x 2/3 = 30%). Also, because not a single student took more than one test in the same year, it is not possible, for example, to determine the correlation between math and Chinese achievement or to evaluate the combined effects of math and Chinese achievement in the same year on corresponding measures of math and Chinese self-concepts. Hence, although extremely valuable for some purposes, these scores are not easily incorporated into the present investigation. We did, however, pursue supplemental analyses in which we used the LISREL procedure for imputing missing values (Joreskog & Sorbom, 1993). Using this procedure, we used T0 English, math, and Chinese achievement (for which we had complete data) to impute scores for missing values at T1. We then used T1 scores (many of which were imputed) to impute scores for missing values at T2, and then imputed T3 scores from the T2 scores. Because data at each year are missing completely at random by design, it may be reasonable to impute missing values as we have done. However, because we needed to impute so many missing values (approximately 2/3 of the scores for each variable were imputed), we only used this data for supplemental analyses and interpreted the results with due caution in relation to findings that did not use these largely imputed test scores from T1, T2, and T3.

Academic Self-concept Measures. The self-concept measures were based on a Chinese version of the SDQ-II (Marsh, 1990b). The original SDQ-II measures eleven areas of self-concept, with three academic domains (English, Math, General Academic), seven nonacademic domains, and a global general self scale (Marsh, 1990b). In the present investigation, a Chinese self-concept scale was created that was strictly parallel with the English self-concept scale (i.e., the wording of the items was the same except for the words English and Chinese). The distinction of English and Chinese self-concepts was necessary to reflect the bilingual emphasis in the Hong Kong school curriculum, but also provides an important basis for evaluating and extending tests of the I/E model. Responses to all the self-concept items were on a 6-point scale ranging from 1 (false) to 6 (true) and administration procedures were based upon those recommended in the SDQ-II test manual (Marsh, 1990b). Responses to this instrument were collected at T2 (8^{th} grade), T3 (9^{th} grade), and T4 (10^{th} grade). Because the focus of the present investigation is on subject-specific measures of academic self-concept as embodied in the I/E model, only responses to the Math, English, and Chinese scales are considered. Psychometric responses to this Chinese version of the SDQ-II were found to be very strong and supported results based on the English version of the instrument (Marsh, 1990b). More specifically, coefficient alpha estimates of reliability were good for all scales (alphas .73 to 94; median = .85) and particularly the academic scales used in this study (math, .94; English, .91; Chinese, .87). Also, confirmatory factor analyses indicated strong support for the a priori factor structure, clearly identifying all 11 scales in the English version of the instrument as well as the Chinese scale that was developed for the present investigation.

Statistical Analysis.

Structural equation models (SEMs) were conducted with LISREL 8 (Joreskog & Sorbom, 1993) using maximum likelihood estimation. A detailed presentation of the conduct of CFA is beyond the scope of the present investigation and is available elsewhere (e.g., Bollen, 1989; Byrne, 1998; Joreskog & Sorbom, 1993; Maruyama, 1998; Schumacker & Lomax, 1996). Following Marsh, Balla, and Hau (1996), and Marsh, Balla, and McDonald (1988) we emphasize the Tucker-Lewis index (TLI), the relative noncentrality index (RNI), and root mean square error of approximation (RMSEA) to evaluate goodness of fit, but also present the χ^2 test statistic and an evaluation of parameter estimates. The TLI and RNI vary along a 0-to-1 continuum in which values greater than .9 are typically taken to reflect an acceptable fit, whereas RMSEAs of less than .05 provide a guideline for an acceptable fit (see Marsh et al., 1996; Schumacker & Lomax, 1996). The RNI contains no penalty for a lack of parsimony so that the addition of new parameters leads to an improved fit that may reflect capitalization on chance, whereas the TLI and RMSEA contains a penalty for a lack of parsimony.

As in other SDQ research and recommended in the test manual (e.g., Marsh, 1988; 1990b; Marsh & Hocevar, 1985), factor analyses were conducted on item-pair scores (or parcels) in which the first two items in each scale are averaged to form the first item pair, the next two items are used to form the second pair, and so forth. Thus, the 10 items in each of the 9 academic self-concept scales (math, English, and Chinese administered on each of three occasions) were used to form 5 indicators that were the basis of subsequent analyses. Analysis of 45 item pairs instead of 90 individual items (10 items/scale x 3 scales x 3 waves of data) is conducted because the responses to item pairs tend to be more reliable, to be more normally distributed, and to have less idiosyncratic variance than do individual items (see Marsh & O'Neill, 1984). Since the T0 achievement scores in each subject are inferred from single indicator measures, an estimated standardized error variance was set to be .10 for each achievement score, resulting in an estimated reliability of .90 for the achievement score (see Joreskog & Sorbom, 1993).

As is inevitable in all large-scale longitudinal studies, many of the students had missing self-concept responses on one or more of the three testing occasions (there were no missing responses for the base year T0 achievement test scores). For the total sample of 9,482, 6,126 (65%) had data for all three occasions, 2,533 (27%) had data for T2 (8^{th} grade) and T3 (9^{th} grade), 724 (8%) had data for T2 only, and 99 (1%) had data for T2 and T4 (10^{th} grade) only. Alternative procedures for accommodating this missing data are presented. First, we conducted analyses using the traditional pairwise deletion for missing data. In this set of analyses, we evaluated the I/E model separately for self-concept responses at T2, T3, and T4 (and T0 achievement scores) and then in a longitudinal model incorporating T2, T3, and T4 responses. There are, however, well known problems in the use of pairwise deletion for missing data – particularly in longitudinal data where substantial amounts of missing data are typical. Hence, we also used the similar response pattern imputation for missing data as operationalized in LISREL 8 (Joreskog & Sorbom, 1993) to impute scores for each of the missing values. Missing self-concept responses were imputed only from scores in the same self-concept domain (e.g., missing math self-concept scores were imputed from other math scores). Because there are well-known potential problems with the use of pairwise deletion for missing data, we focus our presentation on these results with imputed values, but demonstrate that critical parameter estimates for testing the I/E model are very similar for analyses based on pairwise deletion of missing data and the imputed values (e.g., see subsequent discussion of results in Table 1).

In most applications of CFA, a priori models assume that the residual variance (uniqueness plus random error, hereafter referred to as uniquenesses) associated with each measured variable is independent of residual variances associated with other measured variables. However, when the same items are administered to the same participants on multiple occasions (e.g., the math self-concept indicators at T2, T3 and T4), it is likely that the uniquenesses associated with the matching measured variables are correlated (hereafter referred to as 'correlated uniquenesses' that can also be interpreted as method factors). Similarly, if two scales consist of items with parallel wording (e.g., I learn things quickly in English classes; I learn things quickly in Chinese classes), then correlations between measurement errors associated with matching English and Chinese indicators are likely to exist. In each case, if there are substantial correlated uniqueness that are not included in the model, then the estimated correlations between the corresponding latent constructs will be positively biased. However, their inclusion in the model provides a test for these correlated uniquenesses and a control for what might otherwise be a positive bias. Thus, Marsh and Hau (1996) argue that such correlated uniquenesses should always be included in the a priori model, though it is also appropriate to evaluate parameter estimates and fit indices for more parsimonious models that exclude these correlated uniquenesses. Following this recommendation, in the present investigation we evaluate models with no correlated uniquenesses, with correlated uniquenesses for items with parallel wording (for the English and Chinese scales), and with correlated uniquenesses for the same items administered on different occasions (for longitudinal models incorporating data from more than one occasion). Because we argue a priori that such correlated uniquenesses should be included and the results to be presented show that their inclusion results in significantly better fits to the data, we focus primarily on models with correlated uniquenesses. As described later, however, their inclusion had almost no effect on parameter estimates that are critical for testing the I/E model (e.g., see results in Table 1).

Results

Tests of the I/E Model For Each Wave (T2, T3, and T4).

We begin by evaluating tests of predictions based on the extended I/E model (Figure 1B) separately for each wave of data (Tables 1 and 2). Our major focus is on data with imputed values for missing data and correlated uniquenesses (Models 2D, 3D, 4D in Tables 1 and 2). The fit (Table 2) of the a priori (extended) I/E model is extremely good for separate analyses of data from T2 (Model M2D, TLI = .968), T3 (Model M3D, TLI = .968), and T4 (Model M4D, TLI = .980). Critical parameter estimates (Table 1) provide strong support for the I/E model in separate analyses of each of the three waves of data.

1. Correlations among the Chinese, English, and Math self-concept scales are very small (correlations of -.07 to .13) and substantially less than correlations among corresponding Chinese, English, and Math achievement scores (correlations of .67 to .79);
2. mathematics achievement has a substantial positive effect on Math self-concept (path coefficients of .63 to .79), but smaller negative effects on English self-concept and Chinese self-concept (path coefficients of -.35 to -.14) ;
3. English achievement has a substantial positive effect on English self-concept (path coefficients of .48 to .62), but smaller negative effects on Math self-concept and Chinese self-concept (path coefficients of -.26 to -.10) self-concepts; and
4. Chinese achievement has a substantial positive effect on Chinese self-concept (path coefficients of .50 to .61), but smaller negative effects on Math self-concept and English self-concept (path coefficients of -.40 to -.06)).

For these path coefficients that are critical for evaluating predictions of the I/E (see Figure 1B), every path coefficient is statistically significant and in predicted direction in each of the separate analyses of self-concept responses from T2, T3, and T4. These results provide very strong support for the extended I/E model and the stability of the effects over time.

Insert Tables 1, 2 about here

Although not emphasized in our discussion, we compared models with and without correlated uniquenesses and we compared two approaches to dealing with missing data (pairwise deletion and imputation). Although each of the models with correlated uniquenesses is able to fit the data significantly better than models without correlated uniquenesses (e.g., Models M2B vs. M2A, Table 2), the critical parameter estimates for testing the I/E model (Table 1) are nearly identical for each pair of corresponding solutions. Models fit to data based on imputation fit somewhat better than for the corresponding models based on pairwise deletion, but again there is clear support for the I/E predictions based on both the pairwise and imputed approaches. In summary, although we argue for the use of imputed data and models with correlated uniquenesses, support for the I/E model does not depend on this decision.

Support for the I/E model also varies somewhat depending upon the time of data collection (T2, T3, or T4 in Table 1). It is important to emphasize that achievement measures were collected shortly prior to the start of high school (T0), whereas self-concept responses were collected in the second, third, and fourth years of high school. Hence, support of the I/E model in the three waves of data provides a test of the stability of the effects of prior achievement on self-concept over time. We expected the pattern of results would support the I/E model for all three data waves but that the sizes of the path coefficients would become smaller over time. However, the sizes of the path coefficients (both matching and non-matching in Table 1) are only slightly smaller at T4 than T2 or T3, and nearly the same at T2 and T3. These results provide strong support for the stability of the I/E effects over time. Indeed, it is surprising that there was not a greater diminution of support for the I/E model over the period considered.

Tests of the I/E Model Across Three Waves (T2, T3, and T4) of Self-concept Data.

In the longitudinal I/E model, we include all three waves of self-concept data. Although our focus is on the model with correlated uniquenesses based on imputed data, we again summarize the fit of models based on the original data (with pairwise deletion for missing data) and models without correlated uniquenesses (see Tables 2 & 3). In this longitudinal model (Model M5H, Table 3), the latent self-concept factors are well defined in that factor loadings are substantial (median factor loading = .78, Table 3) and the goodness of fit is substantial (TLI = .972, Table 2). We now evaluate support for specific predictions based on the I/E model.

Insert Table 3 about here

Path coefficients relating T0 achievement scores to T2 self-concept responses in the longitudinal model (Table 3, Model M5H) are the same as those for analyses of only the T2 data (Table 1, Models M2D, M3D, M4D with imputed values) in that there are no intervening variables. The effects of T0 achievement on T3 self-concept responses in the longitudinal model (Table 3) are the effects of T0 achievement in addition to those that are mediated by T2 self-concept responses. Not surprisingly, these T3 effects are substantially smaller than the corresponding effects when only the T3 (but not the T2) data are considered. Nevertheless, the pattern of I/E effects is still evident in the T3 responses in that the effects of achievements on matching self-concepts (e.g., T0 math achievement on T3 Math self-concept, .25) are significantly positive whereas those effects on non-matching self-concepts (e.g., T0 math achievement on T3 Chinese self-concept, -.14) are either significantly negative (4 of 6 effects) or nonsignificant. Hence, the effects of prior achievement in the three academic subjects on T3 academic self-concepts– even after controlling the effects of T2 self-concepts – still support predictions based on the I/E model. By T4, there are only very small effects of T0 academic achievement beyond the effects mediated by T2 and T3 academic self-concepts. Even these very small effects at T4, however, tend to be consistent with the I/E model. Thus, all the paths from achievement to matching areas of self-concept are significantly positive whereas the paths to non-matching areas of self-concept are either significantly negative (2 of 6 effects) or non-significant.

Effects of Intervening (T1, T2, T3) Achievements on Relations Between T0 Achievement and T4 Self-concepts.

All tests of the I/E model presented thus far are based on the effects of achievement scores collected prior to the start of high school on self-concepts in the second, third, or fourth years of high school. Although there is some small decline in the size of effects at T4 compared to Times 2 and 3, the reduction is surprisingly small. This suggests that achievement is very stable over time or, perhaps, that the T0 measures of achievement are so important (as a primary basis of entry into high schools of choice) that their effect tends to be strong in spite subsequent changes in relative levels of achievement. In order to pursue this issue we conducted supplemental analyses with results of standardized achievement tests collected at T1, T2, and T3. As noted earlier, the matrix sampling design used to administer these tests meant that students only completed a test in one of the three subjects (math, English and Chinese) in any given year. Although we used the imputation procedure to replace missing tests scores (based in part on the complete set of T0 achievement scores), the substantial amount of missing data (approximately 2/3 of the students had missing data for each test for each year) means that results should be interpreted with appropriate caution. Due in part to this problem, we treated achievement scores at T1, T2, and T3 as multiple indicators of achievement levels during the first half of high school so that a majority of the students had at least one indicator for each area of achievement that was not imputed. In order to provide an unambiguous temporal ordering of the variables, we considered T0 achievement (collected prior to high school), T123 achievement (collected during each of the first three years of high school), and T4 academic self-concept (collected in the fourth year of high school)(Model M4E, Table 4). Of particular interest are the stability of achievement from prior to the start of high school (T0) to achievement early in high school (T1, T2, T3), and the relative contribution of T0 achievement and T123 achievement to T4 academic self-concepts.

The fit of the model to the data is very good (TLI = .970; see Model 4E in Table 2) and the factor structure is well defined (Table 4). Whereas the effects of T123 achievement on T4 self-concept are stronger than those of T0 achievement, it is important to emphasize that the patterns of both sets of effects support the I/E model. For each T123 achievement factor, the effect on the matching area of T4 self-concept is positive (path coefficients of .31 to .60) whereas the effects on non-matching areas of self-concept are all significantly negative or nonsignificant (-.35 to -.06). However, for even the T0 achievement factors, the matching effects were all significantly positive (.24 to .31) whereas the nonmatching effects were all significantly negative (5 of 6 effects, -.14 to -.22) or nonsignificant (.08). Particularly for the critical prediction from the I/E model that achievements in one area will have negative effects on self-concept in nonmatching areas, the support for I/E predictions are as strong or stronger for the effects of T0 achievement than for the effects of T123 achievement. Hence, the T0 achievement scores have a surprisingly strong impact on self-concepts four years latter – even after controlling for the effects of achievement during the three intervening years of high school – and the patterns of results clearly support the I/E model.

Insert Table 4 about here

The two sets of achievement scores (T0 and T123) are very stable over time, and most of the variance in T123 achievement scores can be predicted from T0 achievement scores. Thus, for example the residual variance estimates for T123 achievements in Table 4 are only .16, .11, .11 (corresponding to multiple correlations of .92, .94, and .94 for the prediction of T123 math, English and Chinese achievements from the T0 achievement scores). Nevertheless, the effects of T123 achievement, even after controlling for the effects of T0 achievement, are substantial. This implies that there is change in academic achievement during the first three years of high school that is reflected in the T4 self-concepts. What is surprising, perhaps, is that the T0 achievement scores still have such strong direct effects on the T4 self-concepts – beyond their substantial effect that is mediated through T123 achievement.

Discussion and Implications

There exists a very strong and growing body of support for the I/E model. In evaluating this support, it is important to establish the limits of the model's generalizability. Tests of the cross-cultural support for predictions from a theoretical model developed in one culture to another culture provide an important basis for testing this generalizability. Importantly, previous tests of the I/E model have been conducted primarily on Western countries in which the native language is English. In this respect, the results of the present investigation – based on a large, representative sample of Hong Kong students –

represents an important test of the cross-cultural generalizability of predictions based on the I/E model. Because there was such good support for predictions based on the I/E model in each of three years during the middle high school years, the results clearly support the I/E model.

Extension of the I/E Model to Include Native and Nonnative Languages

The most important contribution of our investigation is the extension of the I/E model to include academic self-concepts and achievement scores in native and nonnative languages. Hong Kong provides an ideal setting for testing this extension in that both Chinese (the native language) and English (the nonnative language) are so important in the high school curriculum and – indeed – in Hong Kong society more generally. Although a few studies supporting the I/E model have considered native languages other than English as the basis of Verbal self-concept, the present investigation is apparently the first study specifically designed to juxtapose the long-term effects of self-concepts and achievements in native and nonnative languages.

The results clearly support a priori hypotheses that nonnative language – as well as native language and mathematics – provide an important basis for the formation of self-concepts in specific school subjects. Although math, English, and Chinese achievements were all highly correlated, math, English and Chinese self-concept were all nearly uncorrelated. Thus, the self-concepts are much more distinct than corresponding areas of achievement. There was clear support for the predictions of the original I/E model (based on math and native language). However, there was also support for new predictions based on the logic of the I/E model that were extended to include nonnative as well as native language. Thus, Chinese achievement had positive effects on Chinese self-concept and negative effects on English and math self-concepts, whereas English achievement had positive effects on English self-concept but negative effects on Chinese and math self-concepts. Furthermore, there was clear support for this pattern of results for self-concepts collected on each of three different occasions up to four years after the collection of achievement at T0.

Longitudinal Extension of the I/E Model

The present investigation is important because it is a true longitudinal study based on data collected over a four-year period starting shortly before students began high school. Particularly because the achievement scores were collected (T0) two years prior to the first wave of self-concept data and each of the three waves of self-concept data (T2, T3, T4) was separated by a full year, there is a clearly established temporal ordering of these variables. Although much of the effects of T0 achievement on T3 self-concepts were mediated by T2 self-concepts, the predicted pattern of results was still evident even after controlling the effects of T2 self-concept. Even more surprising, the weak patterns of effects of achievement on T4 self-concepts beyond the effects that were explained in terms of T2 and T3 self-concepts still provided some support for the I/E predictions. Indeed, for separate analyses of each wave of self-concept data, the strength of the I/E effects were nearly the same at T2 and T3 (two and three years after the collection of the achievement data) and were only diminished slightly at T4 (four years after the collection of the achievement data).

Although clearly supportive of the I/E model, the strength of these effects at T3 and even T4 are somewhat surprising. In particular, even though the achievement scores were collected shortly prior to the start of high school, there was still good support for the I/E predictions in analyses of T4 self-concept measures collected in the fourth year of high school. Why were the effects of prior achievement collected before the start of high school still so strong in the fourth year of high school? Part of the explanation, is that academic achievement is a very stable construct. Furthermore, the particular T0 achievement scores considered in this study were very important to students in that they were the basis of determining whether students were able to attend the high schools of their choice. Because this is such a critical rite of passage in the school life of Hong Kong students (and their parents), the scores that are the basis for this decision are likely to be strong indicators of academic ability that have a long-lasting impact on how students feel about themselves.

In order to pursue these speculations, we conducted supplemental analyses that included both T0 achievement that has been our main focus and subsequent achievement during the first three years of high school (T1, T2, and T3) on T4 self-concepts. These results provide insight into why the effects of T0 achievement are so strong even at T4, four years after the start of high school. Math, English, and Chinese achievements are very stable over time in that T123 achievement during the first three years of high school is highly predictable on the basis of T0 achievement collected prior to the start of high school. Because patterns of achievement at T0 are reflective of patterns of achievements in subsequent years, there continues to be good support for predictions based on T0 achievement even in the fourth year of high school. Thus, for example, students who are relatively good at mathematics and not so strong in Chinese at the start of high school tend to maintain this pattern of results. There are, however, significant direct effects of T0 achievement on T4 self-concept beyond the substantial T0 effects that are mediated through subsequent achievement during the first three years of high school. Hence, the data also support our contention that T0 achievement scores – because of the great importance placed on them as the basis of selection into Hong Kong high schools – continue to have a significant influence on self-concept for many years beyond the effects of subsequent achievement. Hence, there is support for both our suggestions about why T0 achievement had such strong effects on self-concept four years latter. The effects are strong because the achievement scores are very stable over time and apparently, because the T0 achievement scores are so important that they continue to have a lasting effect on subsequent self-concept beyond what can be explained in terms of subsequent achievement.

Broader Implications and Directions For Further Research

Our results – and the I/E model more generally – also have a number of interesting implications for classroom teachers. Research and common sense suggest that positive feedback that lacks credibility is likely to be ineffective. Hence, teachers must judiciously seek to provide positive reinforcement that is credible. In order to achieve this goal, however, teachers must be able to gauge accurately student self-concepts in different academic areas. Whereas SDQ research has found that teachers are able to infer students' academic self-concepts with moderate accuracy, their responses reflect primarily student ability and do not incorporate the internal comparison process the is a central feature of the I/E model (e.g., Marsh & Craven, 1997). Thus, when teachers were asked to infer the self-concepts of low-ability students in different academic areas, they inferred their self-concepts to be uniformly low. In contrast to teacher inferences, the internal comparison process implies that even the least able students may have an average or even above average academic self-concept in their best academic subjects even if their skills are below average in that particular subject. Conversely, when asked to infer the self-concepts of academically gifted students, teachers judged them to be high in all academic areas. In contrast, actual student self-concepts were much more differentiated. Even academically gifted students will be relatively poorer in some school subjects and, consistent with the internal comparison process, may have academic self-concepts that are average or below average in these subjects even if their academic skills are above average. Thus, according to the internal comparison process, everyone feels more positively about themselves in some areas and everyone feels less positively in some other areas.

Although clearly supportive of the extension of the I/E model, the results of our study opens up a host of new directions for further research into the juxtaposition of native and nonnative language self-concepts. Particularly in this Hong Kong setting where English is so dominant as the most important nonnative language, it is not surprising that English – as well as Chinese (the native language) and mathematics – play such a critical role in the formation of academic self-concepts. An important test of the generalizability of these findings is to further evaluate this pattern of results in other settings where there is no one nonnative language that is so dominant or, indeed, where language skills in other than the dominant language may not be highly valued. Thus, for example, in some US school settings there are transitional bilingual education programs that have an implicit or explicit policy to "submerge" the native language of young non-English speaking students in the belief that this will facilitate the acquisition of English language skills (e.g.,. see review Willig, 1985). In contrast to this belief, theoretical models of second language acquisition suggest that mastery of a second language is facilitated by the development and maintenance of good skills, positive attitudes, and a sense of identity in relation to both languages (e.g., Lambert, 1992; Cummins & Swain, 1986; Hakuta & McLaughlin, 1996). Lambert in particular posits a social process model in which second-language learning is a function of prior aptitude and motivation for learning a language. A critical feature of Lambert's model is his emphasis on self-identity. To the extent that second language proficiency is not intended to substitute for first language proficiency (as in the Hong Kong setting), the effects on self-identity are predicted to be positive. However, when the second language is intended to replace the first language, it is predicted that the acquisition of both languages will be undermined and may lead to social alienation. Although these researchers do not focus specifically on academic self-concept in particular school subjects as in the I/E model, it is clear that the juxtaposition of native and nonnative language self-concepts, academic self-concept more generally, and, perhaps, global self-esteem might play a critical role in this research.

References

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

Bond, M. H. (1996). Chinese values. In M. H. Bond (Ed.), The Handbook of Chinese psychology (pp.208-226). Hong Kong: Oxford.

Bong, M. (1998). Tests of the internal/external frames of reference model with subject-specific academic self-efficacy and frame-specific academic self-concepts. Journal of Educational Psychology, 90, 102-110.

Byrne, B. M. (1998). Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications and programming. Mahwah, NJ: Erlbaum

Cummins, J. & Swain, M. (1986). Bilingualism in education: Aspects of theory, research and practice. London, Longman.

Diener, E., Diener, M. & Diener, C. (1995). Factors predicting the social well-being of nations. Journal of Personality and Social Psychology, 69, 851-864.

Hakuta, K.; & McLaughlin, B. (1996). Bilingualism and second language learning: Seven tensions that define the research. In D. Berliner & R. Calfee, et al. Handbook of educational psychology. (pp. 603-621). New York: Prentice Hall.

Hau, K. T. & Salili, F. (1991). Structure and semantic differential placement of specific causes: academic causal attributions by Chinese Students in Hong Kong. International Journal of Psychology, 26, 175-193.

Hau, K. T., & Salili, F. (1996) Prediction of academic performance among Chinese students: Effort can compensate for lack of ability. Organizational Behavior and Human Decision Processes, 65, 83-94.

Hofstede, G. (1991). Culture and organisations. London: McGraw-Hill.

Joreskog, K. G., & Sorbom, D. (1993). LISREL 8: Structural equation modeling with the SIMPLIS command language. Chicago: Scientific Software International.

Lambert, W. E. (1992). Challenging established views on social issues: The power and limitations of research. American Psychologist, 47, 533-542

Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. Psychological Review, 98, 224-253.

Marsh, H. W. (1986). Verbal and math self-concepts: An internal/external frame of reference model. American Educational Research Journal, 23, 129-149.

Marsh, H. W. (1988). Self Description Questionnaire: A theoretical and empirical basis for the measurement of multiple dimensions of preadolescent self-concept: A test manual and a research monograph. San Antonio, TX: The Psychological Corporation.

Marsh, H. W. (1990a). A multidimensional, hierarchical self-concept: Theoretical and empirical justification. Educational Psychology Review, 2, 77-172.

Marsh, H. W. (1990b). Self Description Questionnaire (SDQ) II: A theoretical and empirical basis for the measurement of multiple dimensions of adolescent self-concept: An interim test manual and a research monograph. San Antonio, TX: The Psychological Corporation (Republished in 1992, Publication Unit, Faculty of Education, University of Western Sydney, Macarthur.).

Marsh, H. W. (1993). Academic self-concept: Theory measurement and research. In J. Suls (Ed.), Psychological perspectives on the self (Vol. 4, pp. 59-98). Hillsdale, NJ: Erlbaum.

Marsh, H. W., Balla, J. R., & Hau, K. T. (1996). An evaluation of incremental fit indices: A clarification of mathematical and empirical processes. In G. A. Marcoulides & R. E. Schumacker (Ed.), Advanced structural equation modeling techniques (pp. 315-353). Hillsdale, NJ: Erlbaum.

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., Bong, M., & Yeung, A.S. (Nov., 1999). An extension of the internal/external frame of reference model: A response to Bong (1998). Paper presented at the annual conference of the Australian Educational Research Association (Melbourne, VIC, Australia).

Marsh, H. W., Byrne, B. M., & Shavelson, R. (1988). A multifaceted academic self-concept: Its hierarchical structure and its relation to academic achievement. Journal of Educational Psychology, 80, 366-380.

Marsh, H. W., & Craven, R. (1997). Academic self-concept: Beyond the dustbowl. In G. Phye (Ed.), Handbook of classroom assessment: Learning, achievement, and adjustment (pp. 131-198). Orlando, FL : Academic Press.

Marsh, H. W., & Hau, K-T. (1996). Assessing goodness of fit: Is parsimony always desirable? Journal of Experimental Education, 64, 364-390.

Marsh, H. W., & Hocevar, D. (1985). The application of confirmatory factor analysis to the study of self-concept: First and higher order factor structures and their invariance across age groups. Psychological Bulletin, 97, 562-582.

Marsh, H. W., & Köller, O. (1999). Reunification of East and West German School Systems: Longitudinal Multigroup Study of Academic Self-concept, Achievement, and Multiple Frames of Reference. (in review)

Marsh, H. W., & O'Niell, R. (1984). Self Description Questionnaire III: The construct validity of multidimensional self-concept ratings by late adolescents. Journal of Educational Measurement, 21, 153-174.

Marsh, H. W., & Yeung, A. S. (1998). Longitudinal structural equation models of academic self-concept and achievement: Gender differences in the development of math and English constructs. American Educational Research Journal, 35(4), 705-738.

Maruyama, G. M. (1998). Basics of structural equation modeling. Thousand Oaks, CA: Sage.

McFarland, C., & Buehler, R. (1995). Collective self-esteem as a moderator of the frog-pond effect in reactions to performance feedback. Journal of Personality and Social Psychology, 68, 1055-1070.

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

Skaalvik, E. M., & Rankin, R. J. (1995). A test of the internal/external frame of reference model at different levels of math and verbal self-perception. American Educational Research Journal, 35, 161-184.

Willig, A, C. (1985). A Meta-Analysis of Selected Studies on the Effectiveness of Bilingual Education. Review of Educational Research, 55, 269-318.

Triandis, H. C. (1989). The self and social behavior in different cultural contexts. Psychological Review, 96, 506-520.

Yeung, A. S., & Lee, F. L. (1999). Self-concept of high school students in China: Confirmatory factor analysis of longitudinal data. Educational & Psychological Measurement, 59, 431-450.

Figure 1. Predictions based on the traditional I/E model (A) and the new extended I/E model (1B) that incorporates native and nonnative languages (”+” and “–“ refer to the predicted direction of the path coefficients).

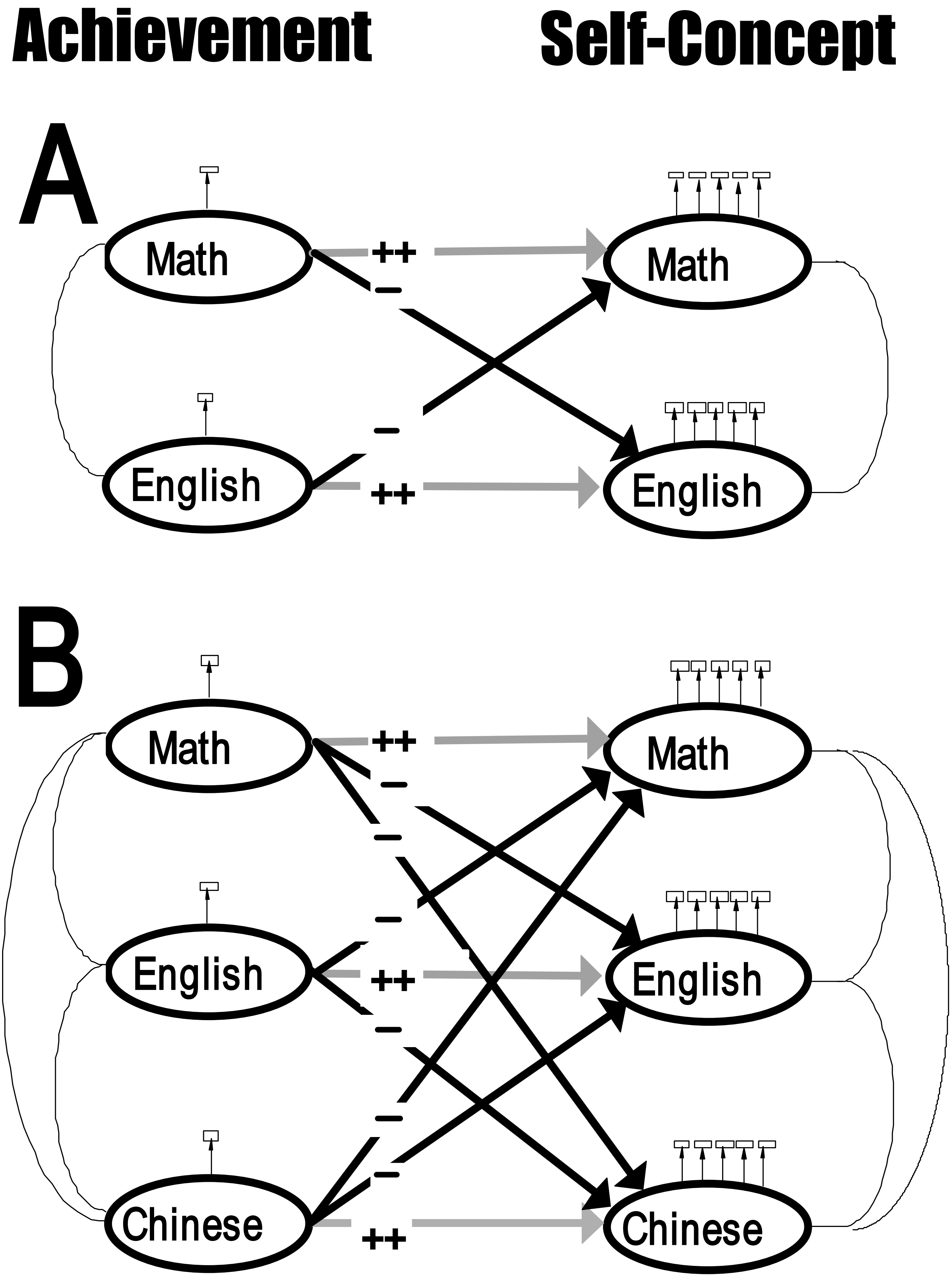

Table 1

Correlations and Path Coefficients Relating Pretest Achievements to Academic Self-concepts. Separate Analyses Based on T2, T3, and T4 Self-concept Responses (using Pairwise Deletion for Missing data or Imputed Missing Values).

		Time 2				Time 3				Time 4			
		Pairwise		Imputed		Pairwise		Imputed		Pairwise		Imputed	
		NoCU	CU	NoCU	CU	NoCU	CU	NoCU	CU	NoCU	CU	NoCU	CU
Variables		M2A	M2B	M2C	M2D	M3A	M3B	M3C	M3D	M4A	M4B	M4C	M4D
Correlations													
Achievement													
MAch	EAch	.67*	.67*	.67*	.67*	.67*	.67*	.67*	.67*	.63*	.63*	.67*	.67*
MAch	CAch	.79*	.79*	.79*	.79*	.79*	.79*	.79*	.79*	.76*	.76*	.79*	.79*
EAch	CAch	.76*	.76*	.76*	.76*	.76*	.76*	.76*	.76*	.72*	.72*	.76*	.76*
Self-concepts													
MSC	ESC	-.03*	-.03*	-.03*	-.03*	.05*	.05*	.05*	.05*	.01	.01	.02	.02
MSC	CSC	-.06*	-.06*	-.06*	-.06*	-.08*	-.08*	-.07*	-.07*	-.04*	-.04*	-.03*	-.03*
ESC	CSC	.12*	.09*	.13*	.08*	.10*	.06*	.10*	.08*	.20*	.17*	.16*	.13*
Path Coefficients													
Matching													
From	To												
MAch	MSC	.79*	.79*	.79*	.79*	.79*	.79*	.77*	.79*	.67*	.67*	.63*	.63*
EAch	ESC	.58*	.58*	.58*	.58*	.62*	.62*	.60*	.62*	.51*	.51*	.48*	.48*
CAch	CSC	.55*	.55*	.55*	.55*	.61*	.61*	.58*	.61*	.53*	.53*	.50*	.50*
Non-Matching													
From	To												
MAch	ESC	-.22*	-.22*	-.22*	-.22*	-.21*	-.21*	-.19*	-.21*	-.18*	-.18*	-.14*	-.14*
MAch	CSC	-.34*	-.34*	-.34*	-.34*	-.35*	-.35*	-.34*	-.35*	-.27*	-.27*	-.27*	-.27*
EAch	MSC	-.17*	-.17*	-.17*	-.17*	-.13*	-.13*	-.12*	-.13*	-.11*	-.11*	-.10*	-.10*
EAch	CSC	-.24*	-.24*	-.24*	-.24*	-.26*	-.26*	-.24*	-.26*	-.22*	-.22*	-.20*	-.20*
CAch	MSC	-.40*	-.40*	-.40*	-.40*	-.40*	-.40*	-.39*	-.40*	-.37*	-.37*	-.33*	-.33*
CAch	ESC	-.14*	-.14*	-.14*	-.14*	-.11*	-.10*	-.11*	-.11*	.01	.01	-.06*	-.06*

Note: M = math, E = English, C = Chinese, Ach = Achievement, SC = Self-concept, NoCU = models with no correlated uniqueness, CU = models with correlated uniquenesses. Separate structural equation models were conducted for self-concept responses from T2, T3, and T4 based on pairwise deletion for missing data and imputed missing values. Presented here are correlations among achievement scores, correlations among self-concept scores, path coefficients leading from achievement to matching areas of self-concept (i.e., math achievement to math self-concept), and path coefficients leading from achievement to non-matching areas of self-concept (i.e., math achievement to English self-concept). Goodness of fit is summarized in Table 2 for each of these models (i.e., M2a, M2B, … M4E).

* $p < .05$

Table 2

Goodness of Fit Indices for Models Varying in: Occasions (T2, T3, T4, or T234; Score Type (Original or Imputed); and Inclusion of A Priori Correlated Uniquenesses

MODEL	CHISQ	DF	RNI	TLI	RMSEA	Correlated Uniquenesses Included
Time 2 Self-concept						
2A Pairwise	3391.37	123	.955	.944	.055	None
2B Pairwise	1906.22	118	.975	.968	.041	Same wording
2C Imputed	3510.15	123	.956	.945	.054	None
2D Imputed	2000.55	118	.976	.968	.041	Same wording
Time 3 Self-concept						
3A Pairwise	4768.17	123	.946	.933	.066	None
3B Pairwise	2502.73	118	.972	.964	.048	Same wording
3C Imputed	4427.44	123	.954	.942	.061	None
3D Imputed	2404.57	118	.975	.968	.045	Same wording
Time 3 Self-concept						
4A Pairwise	3955.76	123	.944	.930	.071	None
4B Pairwise	1855.42	118	.974	.967	.049	Same wording
4C Imputed	2985.10	123	.971	.964	.050	None
4D Imputed	1655.20	118	.984	.980	.037	Same wording
4E[a] Imputed	4438.11	286	.976	.970	.039	Same Wording
T2, T3, T4 Self-concept (Longitudinal Model)						
5A Pairwise	2991.41	1017	.932	.924	.044	None
5B Pairwise	8394.64	972	.958	.951	.035	Same Wording
5C Pairwise	0119.97	972	.948	.940	.039	Longitudinal
5D Pairwise	5636.94	927	.973	.967	.029	Longitudinal & Same Wording
5E Imputed	5985.67	1017	.944	.938	.039	None
5F Imputed	1067.62	972	.962	.956	.033	Same Wording
5G Imputed	1912.85	972	.959	.953	.035	Longitudinal
5H Imputed	7104.49	927	.977	.972	.026	Longitudinal & Same Wording

Note. df = degrees of freedom, RNI = relative noncentrality index, TLI = Tucker-Lewis index. The basic Internal/External model (Figure 1) was fit separately to data from T2, T3, and T4 and then fit to a longitudinal model that included all three waves (T2, T3, T4). For each model, data were based on original data with pairwise deletion for missing data (original) or data in which missing data was imputed (imputed) using the Joreskog and Sorbom approach operationalized in LISREL 8. In different models, correlated uniquenesses were or were not included for items having parallel wording (same wording) of the English and Chinese self-concept scales and for the same items administered on different occasions (longitudinal).

[a] In Model 4E, achievement data from T1, T2, and T3 was included in supplemental analyses to determine how much of the effect of prior (T0) achievement was mediated by subsequent (T1, T2, T3) achievement (see Table 4)

Table 3

Model 5H Relating T0 Achievement to T2, T3, and T4 Self-concepts

	Twelve Latent Factors											
	T0 Achievement			T2 Self-concept			T3 Self-concept			T4 Self-concept		
	MAch	EAch	CAch	MSC2	ESC2	CSC2	MSC3	ESC3	CSC3	MSC4	ESC4	CSC4
Factor Loadings												
Indicator												
1[a]	.95	.95	.95	.78	.82	.76	.83	.85	.79	.83	.85	.79
2				.77*	.78*	.75*	.81*	.81*	.79*	.83*	.83*	.79*
3				.80*	.63*	.61*	.82*	.69*	.66*	.87*	.74*	.68*
4				.79*	.73*	.69*	.83*	.76*	.67*	.85*	.76*	.69*
5				.84*	.65*	.53*	.87*	.71*	.61*	.88*	.75*	.63*
Uniquenesses												
1	.10	.10	.10	.39*	.33*	.42*	.32*	.28*	.38*	.31*	.27*	.38*
2				.41*	.39*	.44*	.34*	.35*	.38*	.31*	.31*	.37*
3				.37*	.60*	.63*	.32*	.52*	.57*	.24*	.45*	.53*
4				.37*	.47*	.53*	.31*	.42*	.55*	.28*	.43*	.52*
5				.30*	.58*	.71*	.24*	.49*	.62*	.23*	.44*	.60*
Path Coefficients												
T0 MAch	0	0	0	0	0	0	0	0	0	0	0	0
EAch	0	0	0	0	0	0	0	0	0	0	0	0
CAch	0	0	0	0	0	0	0	0	0	0	0	0
T1 MSC2	.79*	-.17*	-.40	0	0	0	0	0	0	0	0	0
ESC2	-.22*	.58*	-.14	0	0	0	0	0	0	0	0	0
CSC2	-.34*	-.24*	.56	0	0	0	0	0	0	0	0	0
T2 MSC3	.25*	-.01	-.13*	.67*	.01	.01	0	0	0	0	0	0
ESC3	-.09*	.23*	-.01	.06*	.66*	.02	0	0	0	0	0	0
CSC3	-.14*	-.10*	.26*	-.00	-.01	.58*	0	0	0	0	0	0
T3 MSC4	.05*	.01	-.04	.39*	-.01	.00	.33*	.01	-.02	0	0	0
ESC4	.02	.05*	.02	.03	.35*	.01	-.04*	.38*	.01	0	0	0
CSC4	-.06*	-.08*	.16*	-.02	.04*	.28*	.03	.01	.32*	0	0	0
Variance/Covariance												
T0 MAch	1.00*											
EAch	.67*	1.00*										
CAch	.79*	.76*	1.00*									
T1 MSC2	0	0	0	.76*								
ESC2	0	0	0	.03*	.84*							
CSC2	0	0	0	.02	.12*	.91*						
T2 MSC3	0	0	0	0	0	0	.43*					
ESC3	0	0	0	0	0	0	.02*	.43*				
CSC3	0	0	0	0	0	0	.00	.04*	.59*			
T3 MSC4	0	0	0	0	0	0	0	0	0	.52*		
ESC4	0	0	0	0	0	0	0	0	0	.01	.50*	
CSC4	0	0	0	0	0	0	0	0	0	.02*	.07*	.65*

Note: M = math, E = English, C = Chinese, Ach = Achievement, SC = Self-concept. T0 = pretest; T2 = 8th grade, T3 = 9th grade, T4 = 10th grade. All parameter estimates are presented in completely standardized form (Joreskog & Sorbom, 1993). There are five indicators for each of the three self-concept latent factors representing T1, T3, and T4. There is, however, only a single indictor representing each achievement test score . Because reliability estimates could not be estimated for the achievement factors, a fixed estimate of reliability = .90 (standardized factor loading .95, uniqueness = .10) was used.

[a] The first indicator of each factor were fixed at 1.0 in the unstandardized metric and, thus, were not evaluated in terms of statistical significance.

$p < .05$

Table 4

Model 4E Relating Achievement Prior to the Start of High School (T0) and During the First Three Years of High School (T1, T2, T3) to Academic Self-concept in the Fourth Year of High School (T4)

	Nine Latent Factors								
	T0 Achievements			T123 Achievement			T4 Self-concept		
	MAch	EAch	CAch	MSC2	ESC2	CSC2	MSC3	ESC3	CSC3
Factor Loadings									
Indicator									
1[a]	.95	.95	.95	.77	.89	.86	.83	.85	.79
2				.77*	.85*	.78*	.83*	.83*	.80*
3				.73*	.80*	.74*	.87*	.75*	.68*
4							.85*	.76*	.69*
5							.88*	.75*	.63*
Uniquenesses									
1	.10	.10	.10	.40*	.20*	.26*	.31*	.28*	.38*
2				.41*	.27*	.39*	.31*	.31*	.37*
3				.41*	.36*	.45*	.24*	.44*	.53*
4							.28*	.43*	.52*
5							.23*	.44*	.60*
Path Coefficients									
T0 MAch	0	0	0	0	0	0	0	0	0
T0 EAch	0	0	0	0	0	0	0	0	0
T0 CAch	0	0	0	0	0	0	0	0	0
T123 MAch	.61*	.38*	.00	0	0	0	0	0	0
T123 EAch	.03*	.74*	.22*	0	0	0	0	0	0
T123 CAch	-.09*	.37*	.70*	0	0	0	0	0	0
T4 MSC	.24*	-.22*	-.20*	.60*	-.05	-.18*	0	0	0
T4 ESC	-.14*	.28*	.08	-.08	.48*	-.35*	0	0	0
T4 CSC	-.20*	-.18*	.31*	-.06	-.15*	.31*	0	0	0
Residual Var/Cov									
T0 MAch	1.00*								
T0 EAch	.67*	1.00*							
T0 CAch	.79*	.76*	1.00*						
T123 MAch	0	0	0	.16*					
T123 EAch	0	0	0	.07*	.11*				
T123 CAch	0	0	0	.07*	.05*	.11*			
T4 MSC	0	0	0	0	0	0	.80*		
T4 ESC	0	0	0	0	0	0	.03*	.85*	
T4 CSC	0	0	0	0	0	0	.02	.15*	.92*

Note: M = math, E = English, C = Chinese, Ach = Achievement, SC = Self-concept. T0 = pretest; T1 = 7th grade, T2 = 8th grade, T3 = 9th grade, T4 = 10th grade. All parameter estimates are presented in completely standardized form (Joreskog & Sorbom, 1993). There are five indicators for each of the three self-concept latent factors and three indicators for each T123 achievement factor. Because there was only a single indictor representing each T0 achievement test score, a fixed estimate of reliability = .90 (standardized factor loading .95, uniqueness = .10) was used. Coefficients of 0 were fixed (not estimated) in accordance with the a priori design of the model, as was the first factor loading for the first indicator of each factor.

[a] The first indicator of each factor were fixed at 1.0 in the unstandardized metric and, thus, were not evaluated in terms of statistical significance.

$p < .05$.