

# Math Self-Concept and Mathematics Achievement: Examining Gender Variation and Reciprocal Relations among Junior High School Students in Taiwan

Ching-Yi Lee<sup>1</sup>, Hsin-Yi Kung<sup>2\*</sup>

<sup>1</sup> d.School, Feng Chia University, TAIWAN

<sup>2</sup> Graduate Institute of Education, National Changhua University of Education, TAIWAN

Received 26 November 2016 • Revised 6 November 2017 • Accepted 17 December 2017

## ABSTRACT

The study explored Taiwanese students' mathematics learning at the junior high school level. Utilizing structural equation modeling, it examined the relationships between math self-concept and mathematics achievement with longitudinal data. Participants included 1,256 Taiwanese seventh graders in the first wave and declined slightly to 1,211 eighth graders in the second wave. Findings indicated the following. First, the longitudinal effects were all significant: (a) prior mathematics achievement significantly predicted subsequent math self-concept (skill development model), (b) prior math self-concept significantly predicted subsequent mathematics achievement (self-enhancement model), and (c) the reciprocal effects model was supported, and the effects of achievement tended to become stronger and more systematic. Second, results showed significant gender variation with respect to math self-concept and mathematics achievement. Boys had significantly higher math self-concept than girls, whereas girls exhibited higher mathematics achievement than boys. The implications of these findings for cultivating students' interest in mathematics learning were discussed.

**Keywords:** gender variation, math self-concept, mathematics achievement, reciprocal relations

## INTRODUCTION

Recent studies consistently indicate group differences in mathematics achievement among children from different cultural backgrounds, and this issue has piqued the interest of many educators and psychologists (House, 2006; Shen, 2005; Tsui, 2007). Specifically, cross-cultural studies have supported the finding that students from Taiwan, China, Hong Kong, and Singapore tend to exhibit higher academic achievement, particularly in mathematics, than their American counterparts of elementary and middle-school age (Foy & Olson, 2009; Mullis, Martin, Foy, & Arora, 2012). For example, based on results from the Trends in International Mathematics and Science Study (TIMSS) 2011 of the International Association for the Evaluation of Educational Achievement, Mullis et al. (2012) reported that Taiwanese fourth and eighth grade students' math performance exceeded the international average; fourth graders ranked 4<sup>th</sup> and eighth graders ranked 3<sup>rd</sup> among all the participating countries.

A number of cross-cultural studies have sought to identify the factors influencing achievement, such as parental involvement factors (Yan & Lin, 2005) and schools' instructional practices (Shen, 2005). Further, studies have examined the influence of social-cognitive factors on school performance (Casey, Nuttall, & Pezaris, 2001; Ercikan, McCreith, & Lapointe, 2005; Leung, 2002), and the results have suggested that certain social cognitive factors are critical in influencing students' mathematics achievement. Among these, a positive self-concept is frequently posited as a variable that facilitates certain desired outcomes, such as academic achievement (Casey et al., 2001; Ercikan et al., 2005; Kung & Lee, 2016; Marsh, 1990; Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005; Marsh & Yeung, 1997). Results generally show a positive relationship between students' mathematics achievement and math self-concept. In addition, the general consensus in related studies is that boys outperform girls in mathematics

### Contribution of this paper to the literature

- This study examined the relationships between math self-concept and mathematics achievement with a longitudinal data, and identified good psychometric properties (e.g., high internal reliability) and a well-defined structure of math self-concept and mathematics achievement.
- The longitudinal effects indicated that the reciprocal effects model was supported, and the effects of achievement (i.e., skill development model) tended to be stronger and more systematic.
- Latent mean analysis showed significant gender variation with respect to math self-concept and mathematics achievement. Boys had significantly higher math self-concept than girls, whereas girls exhibited higher mathematics achievement than boys.

achievement, reinforcing the continuing stereotype that girls lack mathematical ability (Casey et al., 2001; Cvencek, Meltzoff, & Greenwald, 2011). This finding calls for up-to-date information about gender differences in math self-concept and mathematics achievement.

Taiwan serves as an interesting case study because its rapid social and economic transformations during the past decades have influenced gender dynamics within the cultural system (Ho, 2013). Much of the work in the mathematics achievement literature with samples of Taiwanese students is cross-sectional in design and uses TIMSS data, which limits investigators from making interpretations about the direction of the relationship between math self-concept and achievement. Hence, conducting research using a longitudinal design might not only lend further credence to the association between self-concept and achievement, but could also provide evidence for directionality; that is, current self-concept either leads to children's enhanced school performance in the future or vice versa. To gain insight into the causal relationship of math self-concept and mathematics achievement, the present study proposes a longitudinal study design using structural equation models that are able to incorporate math self-concept and achievement simultaneously and identify the different factors influencing students' mathematics achievement. Attention will be focused mainly on the reciprocal effects between math self-concept and mathematics achievement, and gender variation issues. The sample includes seventh grade adolescents in Taiwan, with data collected over two consecutive academic years.

## Math Self-Concept and Mathematics Achievement

A positive self-concept (i.e., the set of beliefs we hold about who we are) is a desirable outcome in many educational settings and is frequently posited as a mediating variable that facilitates other desired outcomes, such as academic achievement (Marsh et al., 2005; Skaalvik & Valås, 1999). Recently, however, researchers have emphasized the need to separate the academic components (e.g., the mathematics or English self-concept) from non-academic components of self-concept. Marsh (1993) cited a considerable number of studies showing that although academic achievement is substantially related to academic self-concept, it is almost unrelated to the global and non-academic components of self-concept for people at elementary and middle-school age levels (Byrne, 1996). Hence, it is necessary to accurately measure the self-concept construct from a multidimensional perspective to identify the academic components that are most important for increasing academic achievement.

Researchers have consistently reported a positive relationship between academic self-concept and academic achievement (Byrne, 1996; Casey et al., 2001; Kung, 2009; Ercikan et al., 2005; Marsh et al., 2005; Marsh & Yeung, 1997; Ross, Scott, & Bruce, 2012; Sarouphim & Chartouny, 2017). For example, Ercikan et al. (2005) indicated that students' mathematics confidence was the strongest predictor of achievement, and students' attitudes toward mathematics were the strongest predictors of participation in advanced mathematics courses. In addition, Marsh et al. (2005) argued in support of a reciprocal effects model in which prior self-concepts influence subsequent achievements, known as the self-enhancement model, and prior achievements affect subsequent self-concepts, known as the skill development model. These two models were originally proposed by Calsyn and Kenny (1977). Support for the self-enhancement model would provide a strong justification for the self-concept enhancement interventions that are explicitly or implicitly incorporated into many educational programs. By contrast, the skill development model suggests that academic self-concept emerges principally as a consequence of academic achievement; therefore, the best way to enhance academic self-concept is to develop stronger academic skills. Marsh and Yeung (1997) suggested using longitudinal data to research these issues, measuring self-concept and achievement on at least two occasions (i.e., a two-wave, two-variable design) with latent variable approaches of structural equation modeling. Although the reciprocal effects model has been acknowledged to describe the relationship between academic achievement and academic self-concept, researchers do not yet know which variable exerts the stronger effect, and studies seem to reveal mixed findings. Moreover, few longitudinal studies have focused on specific subject areas (e.g., mathematics). Thus, more studies are needed to evaluate these effects and clarify the mixed findings.

## Developmental Perspectives on the Reciprocal Effects Model

Marsh et al. (2005) indicated that younger children's understanding of academic self-concepts is minimally related to objective outcomes and changes with age. Such a developmental perspective may explain why longitudinal studies seeking to establish the reciprocal effects model have yielded mixed results (Marsh, 1990; Marsh & Yeung, 1997; Skaalvik & Valås, 1999). Skaalvik and Valås (1999) proposed that in the early school years, students' academic self-concepts are not well established and may undergo several processes of shaping and reshaping. When these academic self-concepts become more established and stable, they might increasingly affect performance and study behavior, which might in turn influence academic achievement. Once self-perceptions are more firmly established, the relationships between self-concept and achievement are likely to become reciprocal (Marsh & Yeung, 1997). During late adolescence, academic self-concept might even cause the self-enhancement model to take priority over the skills development model. Evidence indicates that the correspondence between academic self-concept and academic achievement grows stronger with age, at least throughout high school (Marsh, 1990).

However, cross-sectional studies might not be able to capitalize on more extended longitudinal assessments or more sophisticated statistical models. As mentioned above, Marsh (1990) proposed the following optimal design features in research: (a) measurement of academic self-concept and academic achievement (school performance, standardized test scores, or preferably both) at least twice (i.e., a two-wave study); (b) inference of all latent constructs from multiple indicators; (c) use of a sufficiently large and diverse sample to justify the use of structural equation modeling and the generality of the findings; and (d) data fit to a variety of structural equation models that incorporate measurement error and test for likely residual covariance among measured variables.

Longitudinal studies have proposed reciprocal causality between math self-concept and mathematics achievement in Western youth. However, few studies have examined the reciprocal effects model with non-Western samples. Moreover, students from East Asian countries (such as Taiwan) have been found to have lower math self-concept but higher mathematics achievement than their Western counterparts (Foy & Olson, 2009; Mullis et al., 2012). It is therefore necessary to conduct more longitudinal designs using different samples and methods, such as structural equation modeling, to clarify these relationships. In light of these criteria, the present study proposes that a more realistic methodology for examining the relationship between the self-enhancement and skill development models is a *reciprocal effects model*, in which prior math self-concept affects subsequent mathematics achievement and prior mathematics achievement affects subsequent math self-concept. Assuming this model, a multi-occasion study with a longitudinal design was conducted over two consecutive years in Taiwan.

## Gender Differences Related to Math Self-concept

It is important to examine gender differences related to math self-concept because math self-concept may be significant for understanding achievement variation in gender in mathematics education. Although research has highlighted the importance of self-concept in encouraging female to pursue and excel in mathematics (Butt & Dogar, 2014; You, 2010), math self-concept is often overlooked in studies of gender differences, which more often focus on gender differences in achievement (e.g., Isiksal & Cakiroglu, 2008; Lindberg, Hyde, Petersen, & Linn, 2010). However, gender differences in both mathematics attitudes (i.e., math self-concept) and content knowledge might play key roles in creating gender inequality in future math-related courses and career choices. For example, a reason women remain under-represented in the academic disciplines of science, technology, engineering, and mathematics (STEM) fields, as highlighted by previous research, is a lack of math self-concept (Goldman & Penner, 2014). Research has also indicated that the increasing gender gap in math self-concept might later lead to actual gender differences in mathematics achievement (Casey et al., 2001; Cvencek et al., 2011).

Existing research has found no significant gender differences in the global assessment of self-concept between males and females (Rubie-Davies & Lee, 2013), but has found significant gender differences in specific domains of self-concept, specifically in self-concept for mathematical problem-solving, where males reported higher math self-concept than females (Casey et al., 2001; Good, Rattan, & Dweck, 2012; Kung & Lee, 2016), and beliefs of math self-efficacy and fear of failure, where boys favored larger for self-beliefs than girls (Louis & Mistele, 2012; Ross et al., 2012). These results are consistent with those of Cvencek et al. (2011), who indicated that adolescent girls had a lower math self-concept than boys, consistent with gender stereotypes. In addition, Sullivan (2009) examined students' self-concept, and the results indicated that females had lower math self-concept, whereas males had lower self-concept for academics and language.

By contrast, Nagy, Watt, Eccles, Trautwein, Lüdtke, and Baumert (2010) examined gender differences in math self-concept and reported that gender was not significantly related to self-concept. Nagy et al. (2010) concluded that gender-related differences in math self-concept should be questioned. While results seemed to contradict past research implicating gender stereotypes, the studies used indefinite constructs in measuring self-concept, which limited the possibility of an individual obtaining different results across self-concept domains. Therefore, more

studies are needed to examine and clarify these mixed findings. In particular, as most studies have examined only samples of Western students with cross-sectional design, more studies are needed to examine gender-related differences with samples of students from East Asia in math self-concepts from longitudinal and developmental perspectives and to determine whether boys and girls differ by age.

### **Gender Differences in Mathematics Achievement**

Research on differences between the performance of males and females in mathematics has been well documented (Alkhateeb, 2001; Bassey, Joshua, & Asim, 2011; Kung & Lee, 2016; Lindberg et al., 2010; Louis & Mistele, 2012; Sarouphim & Chartouny, 2017). One of the many reasons gender differences in mathematics achievement has been studied so extensively is the quantity of contradictory evidence. For example, studies examining the relationship between gender and mathematics achievement suggest that boys tend to perform better than girls (Bassey et al., 2011; Butt & Dogar, 2014; Marsh & Yeung, 1997; Ross et al., 2012). Recent studies have shown that males continue to outperform females in mathematics achievement, especially on more difficult items (Ross et al., 2012). However, other evidence suggests either a non-existent or declining gender gap in performance, with gender patterns differing between countries. For example, Sullivan (2009) examined students' achievement and found no gender difference in mathematics achievement. Sarouphim and Chartouny (2017) also reported no significant gender differences in achievement in mathematics.

Cross-national studies have suggested that the gender gap in mathematics performance narrows or even reverses in societies with more gender equality (e.g., Sweden and Iceland), but not in those with more gender inequality (e.g., Turkey and Nigeria) (Bassey et al., 2011; Isiksal & Cakiroglu, 2008; Else-Quest, Hyde, & Linn, 2010; Guiso, Monte, Sapienza, & Zingales, 2008). As a result, research on gender differences in mathematics achievement has reported mixed findings. On the one hand, evidence seems to indicate that the gender gap is reducing, but on the other hand, recent research reveals differences in the mathematics performance of boys and girls. As Taiwan has experienced a shift in gender roles due to sociopolitical and economic transformations (Ho, 2013; Ho, Chen, & Kung, 2008), gender-based differences in mathematics performance should become less marked. However, it is unclear whether the gender differences in the mathematics achievement of Taiwanese adolescents correspond to progressive social transformations. Furthermore, comparatively less is known about gender differences in math self-concept among early adolescents. To address this gap in the literature, the present study examines seventh grade students' perspectives on their math self-concept and mathematics achievement to identify any gender differences.

### **Research Questions**

Utilizing the structural equation modeling approach, the present study used longitudinal data to examine the directional effects between math self-concept and mathematics achievement in Taiwan. The study addresses the following questions:

- (1) What is the influence of math self-concept on promoting students' mathematics achievement in Taiwan?
- (2) Is the reciprocal effects model with a longitudinal design (math self-concept vs. mathematics achievement), which is used in this study, appropriate for Taiwanese students, and does the model obtain a reasonable model fit?
- (3) Are there significant gender differences with respect to math self-concept and mathematics achievement?

## **METHODS**

### **Participants**

This study was based on a longitudinal project on public junior high school students' mathematics achievement, in which data were collected in consecutive academic years from representative samples in Taiwan. A multistage sampling process was used to obtain a representative sample of participants and the stratifications were regions (northern, central, and southern), geographic areas (urban and rural settings) within regions, and schools within geographic areas. The northern region included Taipei City, New North City, Taoyuan County, and Hsinchu County; the central region included Miaoli County, Taichung City, Changhua County, Nantou County, and Yunlin County; the southern region included Chiayi County, Tainan City, Kaohsiung City, and Pingtung County. According to the Department of Statistics of Ministry of Education (2016) in Taiwan, the ratio of students in the northern, central, and southern regions was 3:2:2. The second stratified sampling was according to the geographic areas, and the ratio of urban and rural settings was 6:4. With the assistance of city and county governments' bureaus of education, principals, and administrative staff, schools were selected and recommended from the respective educational authorities and researchers. The criteria considered were gender (approximately even numbers of boys

and girls), the school sizes (e.g., large, middle, and small sizes), the average rate of enrolling senior high schools (e.g., lower-performing, average-performing, and better-performing schools), and socioeconomic status (e.g., low, middle, and high socioeconomic status). Once the schools were identified, the classrooms within selected schools could be listed and sampled. To ensure each classroom had an equal chance of being selected, we used a computer program to number and randomly select classrooms to ensure representative participants of individuals. Each selected student was asked to participate in the study by signing a consent form with their parents, homeroom teacher, and the school principal. Before completing the anonymous questionnaire, students were given a brief explanation of appropriate response procedures.

The participants included 1,256 seventh graders in the first wave, which slightly declined to 1,211 eighth graders in the second wave. Over 95% of the students in each selected classroom participated in the study. The regional distribution of the samples was 553 participants from the northern region (44.03%), 360 participants from the central region (28.66%), and 343 participants from the southern region (27.31%). The urban/rural distribution of the samples was 774 participants from urban areas (61.62%) and 482 participants from rural areas (38.36%). The overall sample consisted of 653 boys (51.99%) and 603 girls (48.01%) in seventh grade, and 628 boys (51.86%) and 583 girls (48.14%) in eighth grade.

## Instruments

Data were primarily collected using a questionnaire, with items designed to measure math self-concept in the first and second consecutive years; math achievement levels were also measured in both waves. All the instruments were translated into Chinese by native speakers and back-translated into English for translation verification. This process was repeated until the back-translation into English was deemed to be sufficiently accurate. The students were asked to rate the degree of truth or falseness of each statement: "Overall, how truly or falsely do you agree the following statements?" with bipolar scaling response options on a six-point Likert scale, ranging from 1 (*extremely false*) to 6 (*extremely true*). The bipolar response options asked two things: (1) the direction of the statement (i.e., false or true) and (2) the intensity of the assessment (i.e., slightly, somewhat, mostly, or extremely). All questions were answered from students' perspectives. The following statements indicated bases regarding how the items of questionnaires were developed.

### *Math self-concept questionnaire*

The Self-Description Questionnaire II (SDQ II) developed by Marsh (1988) was adapted for the present study and administered in the middle of the second semester of each academic year. There were two waves of math self-concept data (T1-2 and T2-2). The *Mathematics Self-concept Questionnaire* contains thirteen items that assess three major subscales: (a) competence component, (b) affective component, and (c) comparison component. The students were asked to rate the questions on a six-point Likert scale.

The first composite subscale, consisting of six items, assessed the students' perceptions of their mathematical skills and ability. Participants were asked how true they perceived the following illustrative items to be: "I get good grades in mathematics" and "I learn things quickly in mathematics." Four items were selected to create the second composite subscale, investigating the students' interest in mathematics. Participants were asked to indicate how true they perceived the following illustrative items to be: "I enjoy doing work in mathematics" and "I am interested in mathematics." In addition to 10 items from SDQ II that deal with the math self-concept, the third composite subscale included three items to evaluate students' mathematics learning from the viewpoint of social comparison with their classmates. Sample illustrative items of this subscale were "Compared to my classmates, I'm good at mathematics" and "Compared to my schoolmates, I'm good at mathematics."

Confirmatory factor analysis (CFA) was initially conducted to test the construct validity of math self-concept. Three factors of competence, feeling, and comparison were constructed for the model of math self-concept. Although the  $\chi^2$  was significant ( $\chi^2=516.77, p<.05$ ), it was sensitive to the sample size. As indicated by the alternative fit indices, the CFA indicated reasonable fit. The indices of GFI, CFI, TLI, PNFI, RMSEA, SRMR, and CN were .96, .98, .97, .74, .065, .027, and 329, respectively. Applying Cronbach's alpha, estimates of reliability coefficients for the three subscales and the *Mathematics Self-concept Questionnaire* were, respectively, .80, .85, .92, and .93 for T1-2 and .82, .86, .93, and .94 for T2-2. The data identified good psychometric properties (e.g., high internal reliability) and a well-defined structure (reasonable construct validity) of math self-concept.

### *Mathematics achievement*

Two measurements were used to represent students' mathematics achievement. The first was students' end-of-semester school grades from official records at the end of the first semester of each academic year. Marsh et al. (2005) posited that self-concept should be more strongly related to school grades than to standardized test scores.

They extended this proposal to longitudinal causal modeling studies, suggesting that paths from self-concept to achievement should be stronger for school-based performance measures than for standardized achievement measures (see also Marsh, 1990, 1993). Characteristics such as effort and persistence are likely to have greater impact on examination performance when students are highly motivated to perform well on an examination, such as when these characteristics are an actual part of the grading process, as is typical with school grades. Thus, the effects of prior self-concept on subsequent achievement should be stronger when achievement is based on school grades. However, as school grades varied among schools, it was appropriate to standardize the grades within each class. Therefore, to compare scores across all participants, scaled T scores determining new means and standard deviations within each class were utilized for the analyses rather than raw scores. Thus, every student received a scaled mark within their class that was comparable with other students in other courses and other grades.

The second measurement was teachers' evaluations of students' average mathematics performance in class. The study applied a five-point ratings Likert scale from previous research (e.g., DiPerna & Elliott, 2000; DiPerna, Volpe, & Elliott, 2005), ranging from 1 (*poor/far below grade level expectations*) to 5 (*excellent/far above grade level expectations*), and asked every teacher to evaluate each student's overall performance for items such as class attendance, homework quality, study habits, and fundamental mathematical skills such as measurement, computation, and problem-solving in his/her class. These characteristics were likely to have more impact reflecting students' performance and were actual parts representing students' mathematics achievement. Thus, each teacher completed evaluations for each student in his/her class and evaluated students' overall mathematics performance using statements such as "Please evaluate the student's general performance in mathematics class," using a proficiency rating scale ranging from 1 to 5.

Considering teacher judgement as an indicator of achievement renders teachers as a key source of information about student performance (Martínez, Stecher, and Borko, 2009). Thus, achievement was measured not only by the student's school grades but also by the teachers' judgement of the student's competence. Teachers can assess student achievement with a high degree of accuracy and validity, and because of their sustained interactions with students in the classroom during the school year, teachers gain an understanding of student achievement that is far richer and multidimensional than would be attainable through only standardized tests (Martínez et al., 2009). Further, previous studies have indicated that a teacher's evaluation is a more salient source of feedback that reflects motivational properties likely to be related to students' self-concept (e.g., Kung, 2009; Kung & Lee, 2016). However, as different teachers evaluated each class, and they all had different distributions, the evaluation within each class was standardized so that students were only graded within each course by their corresponding math teacher. To assist in the analysis of standardized evaluation results obtained from these teachers, we scaled the teachers' marks to enable comparisons across different mathematics courses. Because all participants attempt the same mathematics course in grades 7 and 8, we used scaled T scores to determine new means and standard deviations within each class in each grade. Thus, every student received a scaled mark from their teacher for comparisons with students in other courses and grades.

CFA was initially conducted to test the construct validity of mathematics achievement. Although the  $\chi^2$  was significant ( $\chi^2=129.45, p<.05$ ), it was sensitive to the sample size. As indicated by the alternative fit indices, the CFA indicated reasonable fit. The indices of GFI, CFI, TLI, PNFI, RMSEA, SRMR, and CN were .97, .98, .98, .51, .079, .017, and 211, respectively. The reliability coefficients of mathematics achievement, arrived at by utilizing Cronbach's alpha, were .90 and .89 for T1-1 and T2-1, respectively. The data identified good psychometric properties (e.g., high internal reliability) and a well-defined structure (reasonable construct validity) of mathematics achievement.

### Statistical Analysis

A structural equation model-fitting program utilizing the Amos (analysis of moment structure) software package 18.0 was used to conduct the analyses. First, CFA was conducted to test the validity of each latent factor (i.e. mathematics self-concept and mathematics achievement). Second, structural equation modeling was utilized to examine the relationships among these factors. Third, latent mean analysis was conducted to compare gender mean differences in math self-concept. The hypothesized model is shown schematically in [Figure 1](#).

The model evaluation criteria (i.e., the index of assessing the extent to which a model fits an analyzed data set) used to test the fit of the models included the chi-square statistic ( $\chi^2$ ), Goodness-of-fit index (GFI), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Parsimony Normed Fit Index (PNFI), Root-Mean-Square Error of Approximation (RMSEA), Standardized Root Mean Square Residual (SRMR), and Critical N (CN). It should be noted that the  $\chi^2$  statistic is sensitive to sample size; therefore, alternative goodness-of-fit indices were used for the present study. Values of .90 and above for GFI, CFI and TLI were regarded as indicating a reasonable fit (Schumacker & Lomax, 2015). Standardized RMR has an acceptable level when less than .05 and PNFI has an acceptable level when greater than .50 (Schumacker & Lomax, 2015). RMSEA value of 0.05 indicated a close fit and values in the vicinity of .08 indicating a fair fit (Browne & Cudeck, 1993). Hoelter (1983) has suggested that a critical N of 200 or better indicates a satisfactory fit. For the latent mean analysis, a non-significant  $\Delta\chi^2$  statistic indicates

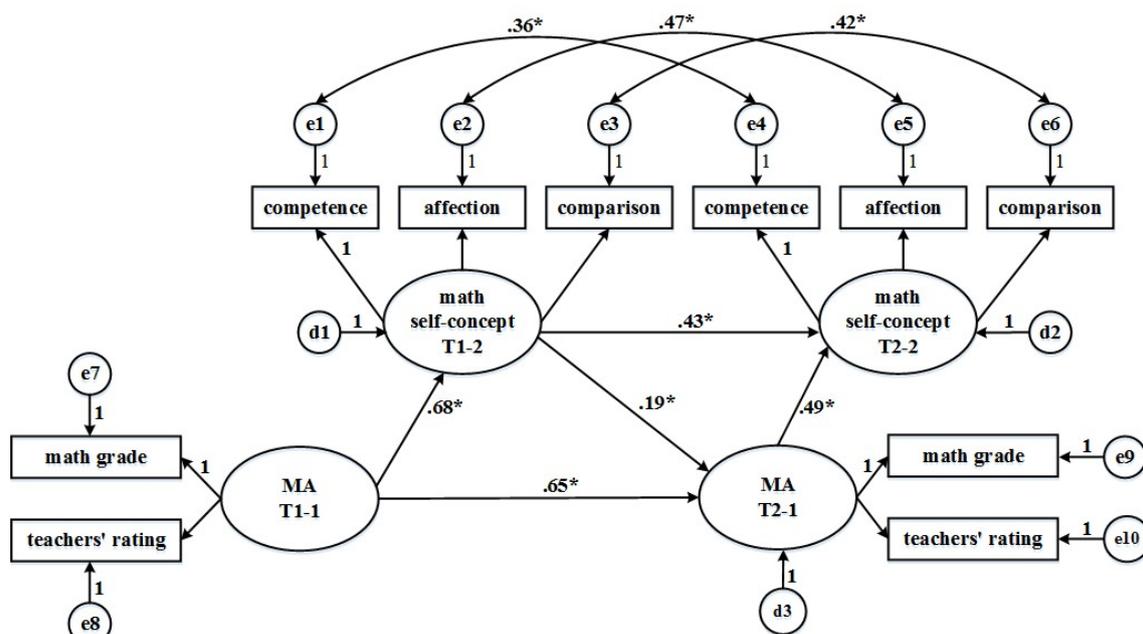


Figure 1. The Hypothesized Model

Table 1. Goodness-of-fit indices for hypothesized model

Goodness-of-fit indices	Hypothesized Model	Decision
$\chi^2$	201.80 ( $df = 28, p = .00$ )	Rejected
GFI	.97	Accepted
CFI	.98	Accepted
TLI	.97	Accepted
PNFI	.59	Accepted
RMSEA (90% CI)	.073 (.072-.074)	Accepted
SRMR	.025	Accepted
CN	276	Accepted

that the compared models were equivalent across groups. In addition, alternative goodness-of-fit indices of  $\Delta CFI$  and  $\Delta RMSEA$  (less than .02) were used to complement the chi-square difference test in comparing the nested models (Fan & Sivo, 2009).

## RESULTS

Preliminary analyses identified no multivariate outliers, and the assumption of normality was never severely violated for any variable, considering the guideline of normality (i.e., skewness < 2; kurtosis < 7) proposed by Curran, West and Finch (1996).

### Confirmatory Factor Analysis

Confirmatory factor analysis was conducted to test the constructs of math self-concept and mathematics achievement. For the factor of math self-concept, the goodness-of-fit indicated that the factor fit the data well. All corresponding factor loadings for the latent factor of math self-concept were significant, ranging from .62 to .92, indicating the moderate to high magnitudes. For the factor of mathematics achievement, the goodness-of-fit also indicated the reasonable fit. All corresponding factor loadings for the latent factor of mathematics achievement were significant, ranging from .86 to .91, also indicating the moderate to high magnitudes.

### Structural Equation Modeling Analysis

The hypothesized model was then evaluated using structural equation modeling to test whether, and to what extent, the model fits the data. The hypothesized model is shown in Figure 1 and the fit index is shown in Table 1.

**Table 2.** Standardized, unstandardized estimate, and S. E. for hypothesized model

	Standardized Estimate	Estimate	S.E.
<b>Math self-concept T1-2</b>			
Competence	.89*	.74	.02
Affection	.66*	.64	.02
Comparison	.93*	1.00 <sup>a</sup>	—
<b>Math self-concept T2-2</b>			
Competence	.91*	.80	.02
Affection	.72*	.68	.02
Comparison	.94*	1.00 <sup>a</sup>	—
<b>MA T1-1</b>			
math grade	.89*	.99	.03
teachers' rating	.91*	1.00 <sup>a</sup>	—
<b>MA T2-1</b>			
math grade	.92*	1.06	.03
teachers' rating	.87*	1.00 <sup>a</sup>	—
Math self-concept T1-2 → Math self-concept T2-2	.43*	.44	.03
MA T1-1 → MA T2-1	.65*	.63	.04
MA T1-1 → Math self-concept T1-2	.68*	.10	.01
MA T2-1 → Math self-concept T2-2	.49*	.08	.01
Math self-concept T1-2 → MA T2-1	.19*	1.18	.22

Note. \*  $p < .05$ .

<sup>a</sup> indicates parameters fixed for identification purpose, not estimated

As indicated by the fit indices, the model fit the data well. Although the  $\chi^2$  was significant, it was sensitive to sample size, and GFI, CFI, TLI, PNFI, RMSEA, SRMR, and CN indicated good fit. The parameter estimates are presented in **Table 2**; all corresponding factor loadings for the all latent factors were significant, with moderate to high magnitudes. The squared multiple correlations were .53 and .55 for math self-concept and mathematics achievement at T2-2 and T2-1, respectively. The composite reliability ( $\rho_c$ ) of math self-concept T1-2, math self-concept T2-2, mathematics achievement T1-1, and mathematics achievement T2-1 were .87, .89, .90, and .80, respectively. The average variance extracted  $\rho_v$  of math self-concept T1-2, math self-concept T2-2, mathematics achievement T1-1, and mathematics achievement T2-1 were .69, .74, .81, and .89, respectively.

The longitudinal effects were all significant. Prior mathematics achievement significantly predicted subsequent math self-concept at T1 and T2 ( $\gamma=.68$  and  $\beta=.49$ , respectively). In addition, prior math self-concept (T1-2) significantly predicted subsequent mathematics achievement (T2-1) ( $\beta=.19$ ). Because these longitudinal paths ( $\gamma$  and  $\beta$ ) were the standardized scores, it is reasonable to compare these paths. Since these paths were comparable, the study found that the magnitudes of the effects from prior mathematics achievement to subsequent math self-concept in two waves (MA T1-1 to math self-concept T1-2 & MA T2-1 to math self-concept T2-2) were consistently greater than the effect from prior math self-concept to subsequent achievement (math self-concept T1-2 to MA T2-1). In other words, the reciprocal effects model was supported and the effects of mathematics achievement tended to be stronger and more systematic.

## Latent Mean Analysis

### *Test of configural invariance*

Hypothesized model was tested between boys and girls. The results were supported a reasonable fit for gender groups (see **Table 3**). Similarly, all manifested items in the model were good indicators of their associated factors for each gender group.

### *Test of metric invariance*

To test for metric invariance, the factor pattern coefficients were constrained to be equal. Thus, the next focus was on testing whether the factor loadings were invariant across the two groups. A  $\chi^2$  difference test was conducted for these nested models. These constraints increased the  $\chi^2$  value, 3.39 with gaining 6 degrees of freedom, and was not statistically significant at  $\alpha=.05$ . **Table 3** presents indices of these nested models, and the values of CFI and RMSEA, generally, were very close for each group. The metric invariance was supported.

**Table 3.** Goodness-of-fit indices for hypothesized Model of gender groups

	$\chi^2$	df	CFI	RMSEA
Baseline model (Model 1)	201.20	54	.985	.049
Factor loading invariance (Model 2)	204.59	60	.985	.046
Structural invariance (Model 3)	227.37	65	.983	.047
Covariance invariance (Model 4)	230.55	68	.983	.047

**Table 4.** The invariance tests for gender groups

	$\Delta\chi^2$	$\Delta df$	$\Delta CFI$	$\Delta RMSEA$	Decision
Model 1 vs. Model 2	3.39	6	.000	.003	Accept
Model 2 vs. Model 3	22.78	5	.002	.001	Accept
Model 3 vs. Model 4	3.18	3	.000	.000	Accept

**Table 5.** Result of latent mean analysis

Construct	Girl group	Boy group	Effect size
Math self-concept T1-2	.00	.50*	.52
Math self-concept T2-2	.00	.34*	.45
Math achievement T1-1	.00	-1.09*	.01
Math achievement T2-1	.00	-1.24*	.01

Note. The latent mean values for girl group were set to zero.

\* indicates that the estimate is significant at  $\alpha=.05$

### Test of scalar invariance

Since the metric invariance has been met, scalar invariance was tested by constraining a series of paths of structural paths and covariance. CFI and RMSEA were also used to evaluate the fit of the model. The values of CFI and RMSEA were close for each group (see [Table 3](#)), and the invariance tests for gender groups were shown in [Table 4](#). Based on the values of  $\Delta CFI$  and  $\Delta RMSEA$  were less than .02 (Fan & Sivo, 2009), structural paths and covariance invariance was supported.

Group differences in the means of latent variables can be estimated only if the latent variables are on the same scale in all groups. Thus, the prerequisites for latent mean analysis are metric invariance and scalar invariance across the multiple groups. With respect to these indices, the invariance test was supported across gender groups for the Taiwanese sample. For identification purposes (Byrne, Shavelson, & Muthén, 1989), the girl group was used as the reference group with its latent mean parameters fixed to zero. Mean parameters for the boy group were free to differ from zero and comparison of the group differences on latent means is based on these differences from zero. [Table 5](#) presents latent mean parameter estimates. Results of latent mean analysis showed significant gender variation with respect to math self-concept and mathematics achievement. Taiwanese boys exhibited higher mean values of math self-concept than the Taiwanese girls in 7<sup>th</sup> and 8<sup>th</sup> grade, and the effect sizes were medium (.52 and .45 for math self-concept of T1-2 and T2-2, respectively). Taiwanese boys exhibited lower mean values of mathematics achievement than the Taiwanese girls in 7<sup>th</sup> and 8<sup>th</sup> grade, however, the effect sizes were very small (.01 and .01 for mathematics achievement of T1-1 and T2-1, respectively).

## DISCUSSION

Mathematics achievement is an important topic in mathematics education, particularly in the field of children's schooling processes. This study examined students' attitudes toward mathematics at length because their effect on learning and performance is significant. Gender difference in mathematics education has been much debated in the literature; yet, few studies have investigated the gender considerations in the relationship between self-concept and achievement while examining reciprocal and longitudinal effects in Taiwan. Furthermore, longitudinal studies examining the stability of math self-concept and gender effects, that is, whether the effect is long-lasting, are rare. The present longitudinal study examined the reciprocal effects model for math self-concept and mathematics achievement, and its relationship to gender differences in Taiwanese adolescents. This paper examined several questions related to self-enhancement and skill development models. Multi-occasion sampling was conducted over two consecutive years to examine gender mean differences significant for math self-concept and mathematics achievement in Taiwan. In summary, the results indicated that (a) prior mathematics achievement significantly predicted subsequent math self-concept (the skill development model); (b) prior math self-concept significantly predicted subsequent mathematics achievement (the self-enhancement model); (c) the reciprocal effects model was supported, and the effects of achievement tended to become stronger and more systematic; and (d) significant

gender variation was found for math self-concept (boys had significantly higher math self-concept than girls) and mathematics achievement (girls had higher mathematics achievement than boys).

### **The Relationships between Math Self-Concept and Math Achievement**

A positive academic self-concept is frequently posited as an important variable for facilitating academic achievement. With respect to the causal ordering of academic self-concept and academic achievement, growing support for the reciprocal effects model would have important implications for the field of education (Marsh, 1990), but less important implications in the field of mathematics education. The present study extended the implications of the well-developed reciprocal effects model, utilizing longitudinal data, with respect to academic self-concept and achievement in mathematics. This study examined whether the model was applicable to math self-concept and mathematics achievement in samples of students from Taiwan. The results were consistent with previous research findings that supported both the skill development and self-enhancement models (Kung & Lee, 2016; Marsh & Martin, 2011; Marsh & Yeung, 1997). Therefore, to promote students' mathematics achievement, it is necessary to determine how this process operates for school students. One approach would be to consider possible intervening variables (e.g., increased effort, enhanced motivation, and persistence in the face of difficulties) that mediate the effect of prior math self-concept on subsequent mathematics achievement (Ercikan et al., 2005). Moreover, since the findings also supported the skill development model, improving students' math problem-solving skills or experience might be an effective approach for reinforcing their math self-concept (Casey et al., 2001).

Notably, this study found that the skill development model had a stronger effect than the self-enhancement model with respect to math self-concept and mathematics achievement. Although the findings reveal that the Taiwanese samples supported both the skill development and self-enhancement models, maintaining students' good performance in mathematics or enhancing students' mathematics skills is an effective way to improve their math self-concept and produce an effect lasting at least two years. These findings highlight the importance of valuing the skill development model in early adolescent years.

### **The Developmental Perspective and the Reciprocal Effects Model**

The developmental perspective holds that the relationship between academic self-concept and academic achievement changes as students' progress through school. This perspective might explain the mixed results of longitudinal studies seeking to establish causal relationships in this area (Byrne, 1996; Marsh, 1990; Marsh & Yeung, 1997; Skaalvik & Valås, 1999). Our findings generally supported the skill development model and the self-enhancement model, and since the samples were from junior-high-school-level students, our study results are consistent with previous findings supporting the reciprocal effects model (Marsh & Yeung, 1997).

The findings were also consistent with the results of Skaalvik and Valås (1999), who advocated a developmental perspective based on the achievement-self-concept relation. Skaalvik and Valås (1999) proposed that in the early elementary school years, students' academic self-concepts are not well established, and self-concept might experience a process of shaping and reshaping, controlled by the influence of academic experience. When the academic self-concept becomes more established and stable, it might increasingly affect performance and study behavior, which might in turn affect academic achievement. Thus, after self-perceptions are more firmly established, the relationship between self-concept and achievement is likely to become reciprocal in early adolescents (Marsh & Martin, 2011). The current study's findings provide important new evidence regarding the generalizability of reciprocal effects in middle school students in Taiwan.

### **Gender Differences Related to Math Self-concept and Mathematics Achievement**

Gender effects were found in this study. A significant gender difference was found for math self-concept, with boys consistently having significantly higher math self-concept than girls. This result supports those of past studies, which indicate that girls give lower evaluations of their math self-concept compared with boys (Casey et al., 2001; Good et al., 2012; Ross et al., 2012; Sullivan, 2009). A possible reason for this finding might be that even in the early adolescent years of academic development, girls are exposed to culturally communicated messages that math is a "boys' subject and not as important for girls" (Guiso et al., 2008). Cvencek et al. (2011) suggested that gender differences for math self-concept in school may stem from an already-established gender bias or inequality emerging in preschool.

Another explanation for this difference might be the power of negative gender stereotypes. Good et al. (2012) indicated that negative stereotypes might carry a strong message that certain groups are less valued or accepted. Schwartz and Sinicrope (2013) proposed that gender bias might cause teachers to expect girls to have negative attitudes toward mathematics and boys to have positive attitudes toward mathematics. Gender stereotypes in mathematics can cause girls in particular to have a lower sense of affinity for mathematics. The present study

speculates that girls who feel less accepted would more likely be less confident in terms of their mathematical abilities than boys who associate themselves with mathematics. Boys with a higher sense of affinity for mathematics will more likely have a more positive math self-concept.

However, for mathematics achievement, the results indicated that the mean levels of mathematics achievement for girls were higher than for boys, but the effect size of the difference is slight. The findings were similar to previous literature (e.g., Alkhateeb, 2001; Kung & Lee, 2016) indicating that females scored higher than males, but with small effect sizes. This finding might be caused by the methods of evaluating students' mathematics achievement. The present study utilized two methods to evaluate students' mathematics achievement. The first method was a scaled T score of students' end-of-semester school test grades measuring mathematics content knowledge from official records at the end of each academic year. Traditionally, studies have shown that girls might do equally as well, or even better, than boys on tests of computation, which require relatively simple cognitive processes, whereas boys tend to perform better on tests requiring more advanced cognitive processing, such as complex problem solving (Lindberg et al., 2010). However, this gender difference in complex problem solving does not emerge until the high school years. The contents of seventh-grade tests might emphasize relatively simple cognitive processes; therefore, girls at this age may perform slightly better than boys. The second method of evaluating students' mathematics achievement was teachers' evaluations on report cards. Here, differences in results might arise because females tend to have higher grades on report cards than do males (Alkhateeb, 2001; Ross et al., 2012), which according to Ross et al. (2012) might be because teachers reward girls with higher evaluations than warranted because of the belief that girls put more effort into mathematics and because girls tend to exhibit comparatively fewer behavioral problems in class than boys.

Recent findings from international studies also suggest that gender differences in mathematics have declined over the years. For example, TIMSS indicated that the differences between girls and boys are minor and have minimal effect sizes (Mullis et al., 2012). Moreover, recent works (e.g., Lindberg et al., 2010; Ross et al., 2012; You, 2010) have shown that girls were consistently receiving better grades than boys in the classrooms because girls are less disruptive and maintain a mastery over their performance goals over time. In conclusion, the results of this study suggest that, while the gender variations in students' math self-concept remain, there was a decline in gender differences in junior high school mathematics achievements in Taiwan. More precisely, the results support that gender-based differences in mathematics performance have become less marked, and the gender gap on achievement appears to be closing.

## CONCLUSION AND IMPLICATIONS

The present study contributed to the study of math self-concept and mathematics achievement in several ways. First, the current investigation is one of few studies to explore the reciprocal effects model using an East Asian sample of Taiwanese students and a longitudinal design. Second, the longitudinal effects indicated that the reciprocal effects model was supported, and the effects of achievement tended to be stronger and more systematic, revealing a developmental progression for reciprocal effects. Finally, latent mean analysis showed significant gender differences with respect to math self-concept and mathematics achievement, suggesting that gender inequality in mathematics requires close attention while Taiwan undergoes social transformations.

The findings of the reciprocal effects model have important implications in the field of mathematics education. Parents, teachers, and school administrators should be aware of these findings and monitor students' mathematics achievement, while considering the role of self-concept. Whereas parents can be reached through newsletters and PTA conferences, teachers and administrators would benefit from more extensive training programs and in-service sessions addressing the importance of incorporating self-confidence into the curriculum. Additionally, the results provide further direction for developing self-confidence strategies for other school subjects, with the possible need to account for students' social class or ethnic background.

Mathematics and science have been stereotyped as male domains (Else-Quest et al., 2010; Lindberg et al., 2010). The gender differences found in this study suggest that stereotypes about female inferiority in mathematics may be prominently projected by children, adolescents, parents, and teachers. Girls with lower math self-concept can prevent themselves from performing at their best and often pursue math-related courses and career paths at lower rates than boys (Goldman & Penner, 2014). To continue to future STEM majors, students (particularly girls) need to have a positive math self-concept; a high aptitude in mathematics is unlikely to lead to a STEM major without a high math self-concept because self-concept provides the positive belief that an individual can succeed in STEM.

Several prescriptive actions can be proposed. Math self-concept begins to form in preschool and early elementary school, with children projecting gender stereotypes about math (e.g., girls displaying lower math self-concepts than boys, and girls being negatively affected by implicit and explicit activation of math-gender stereotypes) (Gunderson, Ramirez, Levine, & Beilock, 2012). To change this dynamic, interventions or educational programs aimed at promoting girls' confidence in mathematics are necessary. In fact, Schwartz and Sinicrope (2013)

found that teacher perceptions and expectations of gender and mathematics change with the completion of a mathematics methods course and classroom experiences. Teachers can make a difference with their actions and overt techniques in the classroom. For example, giving girls more opportunities to participate in mathematics class discussions, inviting girls to join in mathematics-related projects, providing role models of successful women for girls to identify with, and emphasizing the importance of mathematics for future careers and for society in teaching materials and extra-curricular activities are all possible ways to enhance girls' math self-concept.

Other plausible strategies in mathematics education could potentially reduce the confidence gap. Teachers could invest more effort into making mathematics attractive to girls. To overcome the stereotypical belief that mathematics is a male domain, teachers could present role models of female mathematicians, call on girls as frequently as boys to give answers in class, provide equal time for boys and girls to explain their solutions, and create gender-balanced working groups in the classroom. Moreover, school administrators can help teachers recognize that some of their instructional strategies could contribute to stereotyping and thus help students reduce the discrepancies between their self-evaluations and the school standards. Furthermore, schools can train their students to attribute their success to ability and their failure to lack of effort and to respect female students by recognizing their achievements in mathematics.

To summarize, previous research contains surprisingly few non-Western studies employing structural equation models to examine longitudinal data and social cognitive factors pertaining to mathematics learning. This study focused on the relationships between math self-concept and mathematical achievement and, notably, demonstrated the construct validity and reliability of measurements pertaining to the math self-concept constructs in Taiwanese samples. Furthermore, the present study utilized a longitudinal design to more accurately examine the directional effects and gender differences of math self-concept and mathematics achievement. The findings can be used to advance our understanding of the relationships between students' attitudes and achievement and facilitate mathematics performance in students. In addition, these findings can be used to design experimental courses to increase students' math self-concept with particular reference to Taiwanese students' confidence, which falls below the international average in the Index of Students' Self-Confidence in Learning Mathematics (SCM).

Rather than seeking to identify whether one category of students is superior to another, this study aimed to explore how students' self-concept and achievement are related and how gender affects mathematic learning, the results of which might add significant value in seeking ways to enhance teaching and learning for all. From these findings, we suggest that it is important for researchers in the field of mathematics education to become critical educators and, most importantly, be more aware of the social cognitive factors affecting mathematics learning with respect to gender variation.

## ACKNOWLEDGEMENT

This article was supported by research projects funded by the Taiwan Ministry of Science and Technology (NSC 98-2511-S-018-002-M and NSC 99-2511-S-018-014). Our gratitude goes to the anonymous reviewers for their insightful comments and suggestions. We also would like to express our appreciation to all the participants and people assisted in this study.

## REFERENCES

- Alkhateeb, H. M. (2001). Gender differences in mathematics achievement among high school students in the UAE 1991-2000. *School Science and Mathematics, 101*(1), 5-9. doi:10.1111/j.1949-8594.2001.tb18184.x
- Bassey, S. W., Joshua, M. T., & Asim, A. E. (2011). Gender differences and mathematics achievement of rural senior secondary students in Cross River State, Nigeria. *Mathematics Connection, 10*, 56-60.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.
- Butt, I. H., & Dogar, A. H. (2014). Gender disparity in mathematics achievement among the rural and urban high school students in Pakistan. *Pakistan Journal of Social Sciences, 34*(1), 93-100.
- Byrne, B. M. (1996). Academic self-concept: Its structure, measurement, and relation to academic achievement. In B. A. Bracken (Ed.), *Handbook of self-concept: Developmental, social, and clinical considerations* (pp. 287-316). New York, NY: Wiley.
- Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105*(3), 456-466. doi:10.1037/0033-2909.105.3.456
- Calsyn, R. J., & Kenny, D. A. (1977). Self-concept of ability and perceived evaluation of others: Cause or effect of academic achievement? *Journal of Educational Psychology, 69*(2), 136-145. doi:10.1037/0022-0663.69.2.136

- Casey, M. B., Nuttall, R. L., & Pezaris, E. (2001). Spatial-mechanical reasoning skills versus mathematics self-confidence as mediators of gender differences on mathematics subtests using cross-national gender-based items. *Journal for Research in Mathematics Education*, 32(1), 28-57. doi:10.2307/749620
- Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1(1), 16-29. doi:10.1037/1082-989X.1.1.16
- Cvencek, D., Meltzoff, A. N., & Greenwald, A. G. (2011). Math-gender stereotypes in elementary school children. *Child Development*, 82(3), 766-779. doi:10.1111/j.1467-8624.2010.01529.x
- DiPerna, J. C., & Elliott, S. N. (2000). *Academic competence evaluation scales*. San Antonio, TX: The Psychological Corporation.
- DiPerna, J. C., & Volpe, R. J., Elliott, S. N. (2005). A model of academic enablers and mathematics achievement in the elementary grades. *Journal of School Psychology*, 43, 379-392. doi:10.1016/j.jsp.2005.09.002
- Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin*, 136(1), 103-127. doi:10.1037/a0018053
- Ercikan, L., McCreith, T., & Lapointe, V. (2005). Factors associated with mathematics achievement and participation in advanced mathematics courses: An examination of gender differences from an international perspective. *School Science and Mathematics*, 105(1), 5-14. doi:10.1111/j.1949-8594.2005.tb18031.x
- Fan, X., & Sivo, S. A. (2009). Using  $\Delta$ goodness-of-fit indexes in assessing mean structure invariance. *Structural Equation Modeling*, 16, 54-69. doi:10.1080/10705510802561311
- Foy, P., & Olson, J. (2009). *TIMSS 2007 user guide for the international database*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Goldman, A. D., & Penner, A. M. (2014). Exploring international gender differences in mathematics self-concept. *International Journal of Adolescence and Youth*. doi:10.1080/02673843.2013.847850
- Good, C., Rattan, A., & Dweck, C. S. (2012). Why do women opt out? Sense of belonging and women's representation in mathematics. *Journal of Personality and Social Psychology*, 102(4), 700-717. doi:10.1037/a0026659
- Guiso, L., Monte, F., Sapienza, P., & Zingales, L. (2008). Culture, gender, and math. *Science*, 320, 1164-1165, doi:10.1126/science.1154094
- Gunderson, E. A., Ramirez, G., Levine, S. C., & Beilock, S. L. (2012). The role of parents and teachers in the development of gender-related math attitudes. *Sex Roles*, 66(3-4), 153-166. doi:10.1007/s11199-011-9996-2
- Ho, H.-Z. (2013). Gender and Education. In D. Ness, & C.-L. Lin (Eds.), *International education: An encyclopedia of contemporary issues and systems* (pp. 61-63). Armonk, NY: M.E. Sharpe.
- Ho, H.-Z., Chen, W.-W., & Kung, H.-Y. (2008). Taiwan. In I. Epstein, & L. Limage (Eds.), *The Greenwood encyclopedia of children's issues worldwide: Asia and Oceania volume* (pp. 439-464). CT, USA: Greenwood Publishing Group. doi:10.11650/twjm/1500405197
- Hoelter, J. W. (1983). The analysis of covariance structures goodness-of-fit indices. *Sociological Methods & Research*, 11(3), 325-344. doi:10.1177/0049124183011003003
- House, J. D. (2006). Mathematics beliefs and achievement of elementary school students in Japan and the United States: Results from the Third International Mathematics and Science Study. *The Journal of Genetic Psychology*, 167(1), 31-45. doi:10.3200/GNTP.167.1.31-45
- Isiksal, M., & Cakiroglu, E. (2008). Gender differences regarding mathematics achievement: The case of Turkish middle school students. *School Science and Mathematics*, 108(3), 113-120. doi:10.1111/j.1949-8594.2008.tb17814.x
- Kung, H.-Y. (2009). Perception or confidence? Self-concept, self-efficacy and achievement in mathematics: A longitudinal study. *Policy Futures in Education*, 7(4), 387-398. doi:10.2304/pfie.2009.7.4.387
- Kung, H.-Y., & Lee, C.-Y. (2016). The longitudinal reciprocal effects model of junior high school students' mathematics self-concept and mathematics achievement: The perspectives of gender and urban/rural differences. *Chinese Journal of Science Education*, 24(S), 511-536. doi:10.6173/CJSE.2016.24S.04
- Leung, F. K. S. (2002). Behind the high achievement of East Asian students. *Educational Research and Evaluation*, 8(1), 87-108. doi:10.1076/edre.8.1.87.6920
- Lindberg, S. M., Hyde, J. S., Petersen, J. L., & Linn, M. C. (2010). New trends in gender and mathematics performance: A meta-analysis. *Psychological Bulletin*, 136(6), 1123-1135. doi:10.1037/a0021276
- Louis, R. A., & Mistele, J. M. (2012). The differences in scores and self-efficacy by student gender in mathematics and science. *International Journal of Science and Mathematics Education*, 10, 1163-1190. doi:10.1007/s10763-011-9325-9

- Marsh, H. W. (1988). *Self Description Questionnaire I: 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. (1990). Causal ordering of academic self-concept and academic achievement: A multiwave, longitudinal panel analysis. *Journal of Educational Psychology, 82*(4), 646-656. doi:10.1037/0022-0663.82.4.646
- Marsh, H. W. (1993). Academic self-concept: Theory measurement and research. In J. Suls (Ed.), *Psychological perspectives on the self* (pp. 59-98). Hillsdale, NJ: Erlbaum.
- Marsh, H. W., & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. *British Journal of Educational Psychology, 81*, 59-77. doi:10.1348/000709910X503501
- Marsh, H. W., & Yeung, A. S. (1997). Causal effects of academic self-concept on academic achievement: Structural equation models of longitudinal data. *Journal of Educational Psychology, 89*(1), 41-54. doi:10.1037/0022-0663.89.1.41
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects models of causal ordering. *Child Development, 76*(2), 397-416. doi:10.1111/j.1467-8624.2005.00853.x
- Martínez, J. F., Stecher, B., & Borko, H. (2009). Classroom assessment practices, teacher judgments, and student achievement in mathematics: Evidence from the ECLS. *Educational Assessment, 14*, 78-102. doi:10.1080/10627190903039429
- Mullis, I. V., Martin, M. O., Foy, P., & Arora, A. (2012). *TIMSS 2011 international results in mathematics*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Nagy, G., Watt, H. M., Eccles, J. S., Trautwein, U., Lüdtke, O., & Baumert, J. (2010). The development of students' mathematics self-concept in relation to gender: Different countries, different trajectories? *Journal of Research on Adolescence, 20*(2), 482-506. doi:10.1111/j.1532-7795.2010.00644.x
- Ross, J. A., Scott, G., & Bruce, C. D. (2012). The gender confidence gap in fractions knowledge: Gender differences in student belief-achievement relationships. *School Science and Mathematics, 112*(5), 278-288. doi:10.1111/j.1949-8594.2012.00144.x
- Rubie-Davies, C. M., & Lee, K. (2013). Self-concept of students in higher education: are there differences by faculty and gender? *Educational Studies, 39*(1), 56-67. doi:10.1080/03055698.2012.671513
- Sarouphim, K. M., & Chartouny, M. (2017). Mathematics education in Lebanon: Gender differences in attitudes and achievement. *Education Studies in Mathematics, 94*(1), 55-68. doi:10.1007/s10649-016-9712-9
- Schumacker, R. E., & Lomax, R. G. (2015). *A beginner's guide to structural equation modeling* (4<sup>th</sup> ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Schwartz, C. S., & Sinicrope, R. (2013). Prospective elementary teachers' perceptions of gender differences in children's attitudes toward mathematics. *School Science & Mathematics, 113*(8), 410-416. doi:10.1111/ssm.12045
- Shen, C. (2005). How American middle schools differ from schools of five Asian countries: Based on cross-national data from TIMSS 1999. *Educational Research and Evaluation, 11*(2), 179-199. doi:10.1080/13803610500110810
- Skaalvik, E. M., & Valås, H. (1999). Relations among achievement, self-concept, and motivation in mathematics and language arts: A longitudinal study. *The Journal of Experimental Education, 67*(2), 135-149. doi:10.1080/00220979909598349
- Sullivan, A. (2009). Academic self-concept, gender and single-sex schooling. *British Educational Research Journal, 35*(2), 259-288. doi:10.1080/01411920802042960
- The Department of Statistics of Ministry of Education (2016). *Education statistics inquiry network*. Retrieved from <https://stats.moe.gov.tw/>
- Tsui, M. (2007). Gender and mathematics achievement in China and the United States. *Gender Issues, 24*, 1-11. doi:10.1007/s12147-007-9044-2
- Yan, W., & Lin, Q. (2005). Parent involvement and mathematics achievement: Contrast across racial and ethnic groups. *The Journal of Educational Research, 99*(2), 116-127. doi:10.3200/JOER.99.2.116
- You, Z. (2010). Research in brief: Gender differences in mathematics learning. *School Science and Mathematics, 110*(3), 115-117. doi:10.1111/j.1949-8594.2010.00028.x