TSUNG-HAU JEN, CHE-DI LEE, CHIN-LUNG CHIEN, YING-SHAO HSU
and KUAN-MING CHEN

PERCEIVED SOCIAL RELATIONSHIPS AND SCIENCE LEARNING OUTCOMES FOR TAIWANESE EIGHTH GRADERS: STRUCTURAL EQUATION MODELING WITH A COMPLEX SAMPLING CONSIDERATION

Received: 14 February 2012; Accepted: 28 June 2012

ABSTRACT. Based on the Trends in International Mathematics and Science Study 2007 study and a follow-up national survey, data for 3,901 Taiwanese grade 8 students were analyzed using structural equation modeling to confirm a social-relation-based affection-driven model (SRAM). SRAM hypothesized relationships among students' perceived social relationships in science class and affective and cognitive learning outcomes to be examined. Furthermore, the path coefficients of SRAM for high- and low-achieving subgroups were compared. Given the 2-stage stratified clustering design for sampling, jackknife replications were conducted to estimate the sampling errors for all coefficients in SRAM. Results suggested that both perceived teacher–student relationships (PTSR) and perceived peer relationships (PPR) exert significant positive effects on students' self-confidence in learning science (SCS) and on their positive attitude toward science (PATS). These affective learning outcomes (SCS and PATS) were found to play a significant role in mediating the perceived social relationships (PTSR and PPR) and science achievement. Further results regarding the differences in SRAM model fit between high- and low-achieving students are discussed, as are the educational and methodological implications of this study.

KEY WORDS: complex sampling, large-scale survey, learning motivation, science achievement, self-determination theory, social relationships, structural equation modeling, TIMSS

Introduction

Contemporary science education reforms emphasize interactive and constructive learning in which learners learn from the interaction among prior knowledge, concurrent experiences, and human interactions within socio-cultural contexts (United States National Research Council, 2007). However, how learners' perceived social environment and attributes influence learning outcomes is complex and not well documented. Large international assessment data sets provide the basis for important secondary analyses that have implications regarding student, school, and cultural attributes far beyond the league tables about who's in first place flowing from these

International Journal of Science and Mathematics Education (2013) 11: 575–600

surveys (Yore, Anderson & Chiu, 2010). The current study benefits from the comprehensive data set of large-scale surveys to confirm the proposed social-relation-based affection-driven model (SRAM), illustrating the relations among learners' perceived social environment, attributes, and academic achievement and further to differentiate mechanisms of motivation in learning science between low and high achievers. In addition, an important methodological technique of using complex sampling of large-scale surveys is demonstrated to advise researchers of future studies.

Education Context in Taiwan

Asian countries value highly the benefits brought by education and academic degrees. Cultural and family expectations reflect the Confucian philosophy in which children work hard in school and respect the authority of their teachers so as to achieve high academic grades and access to higher education. The Joint Public Senior High School Entrance Examination had been the only way for junior high school students to continue their senior high school education and, therefore, caused much psychological distress on students preparing for the examination. However, in 2002, the Multi-route Promotion Program was implemented to provide various ways and opportunities for entrance to senior high school, but examination pressure still continues among students and their families. Most junior high school students spend significant time studying and even attending cram schools to prepare for the Basic Competency Test, which is an important performance consideration for senior high school entrance. Under this content-oriented test circumstance, most junior high science teachers pay more attention to students' cognitive achievement than to their affective learning outcomes (Tsai & Kuo, 2008). Even though the national curriculum guidelines set the goal of enhancing students' science attitudes and confidence in science learning, Taiwanese grade 8 students were ranked second-worst in confidence and positive attitude toward science learning among countries participating in the 2007 Trends in International Mathematics and Science Study (TIMSS), while performing well above the international average in science knowledge. In addition, the percentages of Taiwanese grade 8 students with high positive attitudes toward science, high values for science, and high self-confidence in learning science decreased from TIMSS 1999 to 2007. The government in Taiwan has funded many research projects to enhance students' self-confidence in learning science and positive attitude toward science. Some of these projects have focused on innovations in school science teaching while others aimed to improve students' motivation to learn science.

Learning Motivation and Classroom Culture

The potential relationships among students' attitudes, beliefs, interest, motivation, and achievement seem reasonable, but the actual mechanisms that connect student traits and achievement are not transparent. Perceived social supports are related to enhancing students' academic self-confidence and attitude (e.g. Nelson & DeBacker, 2008; Wolf & Fraser, 2008). According to the self-determination theory, there are three basic psychological needs essential to intrinsic motivation—relatedness, competence, and autonomy (Ryan & Deci, 2000b). Relatedness is the need to experience supportive social relationships with and connected to others. Competence is the need of being confident in the efficacy of one's abilities. Autonomy is the need to feel autonomous rather than feel compelled in one's actions. Intrinsic motivation, rather than extrinsic demands, enhances one's self-concept and engagement leading to higher achievement. That is, the self-determination theory proposes that beliefs about self play a mediator role/function between the psychological needs and achievement behavior. However, the self-concept differs across cultures (Markus & Kitayama, 2003), and the relationship among the three basic psychological needs is also culture-dependent (Keller, 2012). For example, people from collectivistic Asian cultures could internalize extrinsic demands through fulfilling the need for relatedness because they endorse interdependent self-construal (Bao & Lam, 2008; Ryan & Deci, 2000a). Therefore, for collectivistic cultures, enhancement of learners' positive social relationships should play an essential role in the needs of autonomy and competence and, in turn, improve learning achievement.

Other studies have suggested that students at high- or low-achieving levels differed in their engagement in learning activities. For example, Cosmovici, Idsoe, Bru & Munthe (2009) observed that improvement in the perceptions of learning environment, including perceived emotional and academic supports from teachers or peers, exerts different effects on learning motivation for students at different achievement levels. Ng, Kenney-Benson & Pomerantz (2004) indicated that the effects of a supportive parent–child relationship on children's performance are different between low and high achievers.

Research Problem Context

Given Taiwan's sociocultural context (collectivistic but competitive) and the desire to enhance students' self-confidence in learning science and positive attitude toward science, the current study aimed to illustrate how Taiwanese students' perceived social relationship (PSR) in science class influences their affective and cognitive learning outcomes and to differentiate the mechanisms of learning motivation for students at different achievement levels. The specific research questions that guided this study were:

1. To what extent can perceived social relationships in science class predict students' affective learning outcomes?
2. To what extent can students' affective learning outcomes predict their cognitive learning outcomes?
3. To what extent can perceived social relationships in science class predict students' cognitive learning outcomes through the mediation of students' affective learning outcomes?

So as to agree with the data sets of TIMSS 2007 and its follow-up national survey of Taiwan, the PSR that corresponded to relatedness in the self-determination theory were further separated into perceived teacher–student relationship (PTSR) and perceived peer relationship (PPR) in science class. PTSR refers to students' respect, personal regard, and interpersonal connection for/with their science teacher and perceived support from their science teacher (Hardre, Chen, Huang, Chiang, Jen & Warden, 2006); PPR refers to perceived mutual acceptance and cohesiveness within peer groups in science class (Wolf & Fraser, 2008). Affective learning outcomes include competence and autonomy reflected by students' self-confidence in learning science (SCS) and positive attitude toward science (PATS; Martin & Preuschoff, 2008). SCS indicates a student's self-evaluation of his/her ability in solving science problems or working on scientific activities; PATS indicates the eagerness to engage in activities related to science and learning science. Finally, cognitive learning outcome refers to students' science achievement (SA) in TIMSS 2007. The hypothetical model was formulated on the basis of these variables as below.

Social Relationship and Affective Learning Outcome. The fulfillment of relatedness plays a pivotal role on feeling competent and autonomous not only for people of collectivistic cultures but abundant studies also suggest a relationship among them for other cultures. For example, perceived teacher support correlated positively with academic self-confidence for Australian secondary school students (Dorman, 2001), and perceived peer support and belongingness had a positive impact on the academic self-confidence of American grades 6, 7, and 9 students (Nelson & DeBacker, 2008). Based on data from Taiwanese senior high school students, Hardre et al. (2006) observed that both perceived teacher and peer support correlated with academic self-confidence and that the relationship between perceived peer support and academic self-confidence was stronger than the one between perceived teacher support and academic self-confidence. Other studies indicate the effects of a supportive teacher–student relationship and a cohesive peer relationship on positive attitudes. In Wolf & Fraser's (2008) study, both teachers' support and students' cohesiveness correlated with students' positive attitude toward science, and teachers' support had a greater effect than students' cohesiveness on students' positive attitude toward science.

Affective Learning Outcome and Academic Achievement. Students' affective attributes are important indices of educational effectiveness; therefore, affective attributes are included as items in large-scale educational evaluation studies [e.g. TIMSS and Programme for International Student Assessment (PISA)]. In TIMSS, the affective attributes are reflected directly by SCS and PATS. Since Bandura (1977) proposed the theory of belief-in-self in a specific task domain, academic self-confidence has been recognized as the most important affective factor in student achievement. In their meta-analysis of 36 studies, Multon, Brown & Lent (1991) concluded that academic self-confidence is a significant predictor of academic performance. The relationship between students' science attitude and their learning achievement has been a concern in science education research; however, the strength of that relationship is controversial. For example, Weinburgh's (1995) meta-analysis supported the positive relationship between the two factors, whereas in Shrigley's (1990) review, attitude toward science correlated with learning achievement only when the measurement of attitude was refined.

Hypothetical Model and Expected Results

According to self-determination theory (Ryan & Deci, 2000b) and the previously reported empirical studies, various correlations among students' perceived social support, affective learning outcome, and cognitive learning outcome exist; however, most of the studies merely explored the relationships among some of these variables. The current study proposes a SRAM (Figure 1) to (a) comprehensively consider all relations among these variables and (b) enact the planned secondary analysis procedures. We assume that PTSR and PPR exert direct effects on both affective learning outcome (SCS and PATS) and, in turn, that SCS and PATS influence the cognitive learning outcome (SA). That the PSR do not have direct effect on the

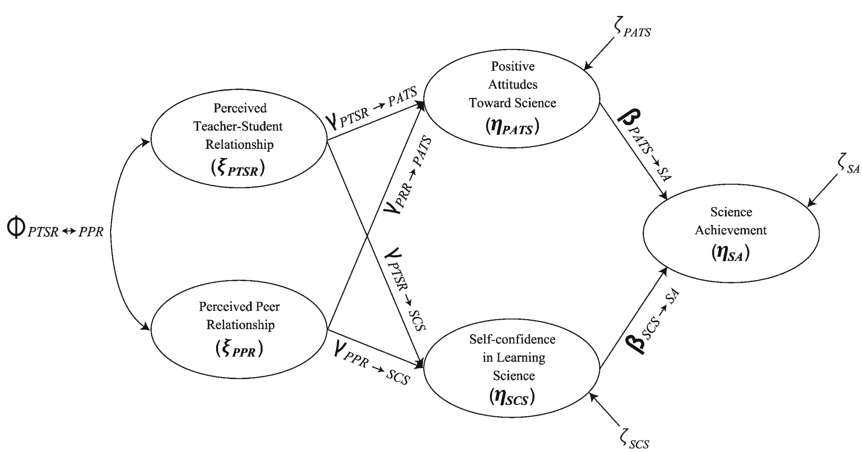

Figure 1. The social-relation-based affection-driven model

cognitive learning outcome in the SRAM are supported by other studies (Erlauer, 2003; Taylor & Lonsdale, 2010).

SRAM can be represented mathematically by a series of simultaneous regression equations as follows:

$$\eta_{\mathrm{SA}} = \beta_{\mathrm{PATS}\rightarrow\mathrm{SA}} \cdot \eta_{\mathrm{PATS}} + \beta_{\mathrm{SCS}\rightarrow\mathrm{SA}} \cdot \eta_{\mathrm{SCS}} + \zeta_{\mathrm{SA}}$$
$$\eta_{\mathrm{PATS}} = \gamma_{\mathrm{PTSR}\rightarrow\mathrm{PATS}} \cdot \xi_{\mathrm{PTSR}} + \gamma_{\mathrm{PPR}\rightarrow\mathrm{PATS}} \cdot \xi_{\mathrm{PPR}} + \zeta_{\mathrm{PATS}}$$
$$\eta_{\mathrm{SCS}} = \gamma_{\mathrm{PTSR}\rightarrow\mathrm{SCS}} \cdot \xi_{\mathrm{PTSR}} + \gamma_{\mathrm{PPR}\rightarrow\mathrm{SCS}} \cdot \xi_{\mathrm{PPR}} + \zeta_{\mathrm{SCS}}$$

The η and ξ indicate the measures of endogenous latent construct and exogenous latent construct, respectively. The residual is denoted as ζ. Their subscripts are the names of the latent constructs. β is the regression coefficient for the endogenous variable whereas γ is for the endogenous variable, and their subscripts specify the direction from the predictor to the predicted variable. In addition, $\phi_{\mathrm{PTSR}\leftrightarrow\mathrm{PPR}}$ (Figure 1) refers to the correlation between the two exogenous latent constructs. SRAM was used to interpret both the total data set and the separate data of high- and low-achieving students.

Therefore, based on our proposed framework and other research results, all the path coefficients in SRAM are expected to be positive, reflecting that the students' perceived social relationships (PTSR and PPR) in science class positively predict their affective learning outcomes (PATS and SCS) and their affective learning outcomes positively predict their cognitive learning outcomes (SA). In addition, students' affective outcomes are expected to mediate the effect of their perceived social relationships in science class on SA. However, these effects might be different between high- and low-achieving students.

METHOD

Structural equation modeling (SEM) was adopted for the secondary analysis on the TIMSS 2007 data set and its follow-up national survey to confirm the proposed SRAM. SEM technique is beneficial for a secondary analysis to "allow measurement error, multiple indicators and test for confounding variables" (Bollen, 1989, p. 73), whereas the traditional regression technique assumes that all variables are measured without error. Traditional multiple regression or path analysis techniques use raw or standardized scores without considering the measurement errors of variables to estimate the regression coefficients; therefore, the regression weight of a predictor is attenuated even though a good reliability of the instrument is reported (McCoach, Black & O'Connell, 2007). From this viewpoint, SEM estimates the effect sizes more accurately than the traditional path analysis method does.

There are two main advantages for using the database of large-scale surveys for a secondary analysis: The scrupulous sampling design assures the sample is representative of the whole population, and the quality of the instrument is confirmed by pilot studies and by content and methodology experts. However, some limitations of the instruments and the sampling method need to be considered. First, items of an instrument are defined and developed to meet the specific purpose of a large-scale survey; therefore, only those items in agreement with the constructs defined for a secondary analysis can be used. This requirement drastically reduces the number of items, but fortunately, enough items can still be identified due to the large item pool. Also, if a secondary analysis aims to examine the relationships among variables at a population level, only a reasonable number of items are needed to produce sensitive and reliable results due to the large sample size. Second, because most large-scale surveys adopt a two-stage stratified cluster sampling, the procedure of variance estimations becomes extremely sophisticated and daunting as compared to a simple random selection at the individual level. In the current study, we used jackknife replications as the standard procedure to estimate the parameter variances (Foy, Galia & Li, 2008).

Finally, we recognize the debate whether a cross-sectional survey can be used to obtain a causal model because it only provides correlations among variables. Although not sufficient, correlation is a necessary condition of causality. The current study adopted a hypothesis–deduction–confirmation approach to confirm, rather than to verify, the theoretical model: SRAM is specified at first and then, based on the causal relations embedded in SRAM, the specific patterns of correlations among variables are confirmed.

Data Source

The data of 4,046 Taiwanese grade 8 students were originally identified as being part of TIMSS 2007 and its concurrent follow-up national survey. Given the unequal probability of a student to be sampled, the 4,046 students' SAs were separated into high-achieving and low-achieving groups according to the weighted median. A listwise deletion of 145 students with missing data resulted in a total final sample of 3,901 participants, including the high-achieving subgroup (HAG; $n_{\mathrm{HAG}} = 1{,}956$) and the low-achieving subgroup (LAG; $n_{\mathrm{LAG}} = 1{,}945$), was retained to validate the SRAM and examine the research questions.

Instruments

Both the student background questionnaire and the released science test scores (plausible values) of TIMSS 2007 and the concurrent follow-up survey were adopted for secondary analysis. The questionnaire included information about attitude toward science and mathematics, self-confidence in learning science and mathematics, perceived classroom activities, and home background (Martin & Preuschoff, 2008). The survey collected information of local interests, such as PSR in the classroom, frequency of attending cram school, and family income. Three panelists (two experts in science education and one expert in psychology) were recruited to screen the item pools according to the definitions of the four constructs (PTSR, PPR, SCS, PATS) and to elaborate the measurement model. They identified 13 items from the pool of 52 items that did meet the agreement criteria. Only consensus items were retained for the confirmatory factor analyses and the following SEMs. All items and their corresponding constructs are provided in Table 1.

The two independent variables, PTSR and PPR, contained three items each measured in the follow-up survey. One of the mediators, SCS, was assessed by four items and the other mediator, PATS, by three items from the questionnaire. All items for the four constructs used a four-point Likert scale ranging from "agree a lot" to "disagree a lot."

The dependent variable, SA, refers to the plausible value (PV) provided in the TIMSS 2007 database. Each item in the TIMSS science test ($N = 210$) has two dimensions. The content dimension specifies the subject matter within science (i.e. physics, chemistry, biology, and earth science), and the cognitive dimension specifies the thinking process (i.e. knowing, applying, reasoning). The 210 items were distributed into 14 booklets of about 30 items each. Item response theory (IRT) was used to equate the scales across the booklets based on the common items between

TABLE 1

Indicators and variable types in SRAM

Construct	*Items (item code)*	*Variable type*
PTSR	I like my science teacher. (PTSR1)[a]	Ordinal
	My science teacher does care about me. (PTSR2)[a]	Ordinal
	When I encounter difficulties, my science teacher will assist me in solving them. (PTSR3)[a]	Ordinal
PPR	I think my peers like to be in a group with me when group activities are conducted. (PPR1)[a]	Ordinal
	I can express my opinions in science class comfortably and need not to worry that my peers will laugh at my ideas. (PPR2)[a]	Ordinal
	I have good communication with my peers in science class. (PPR3)[a]	Ordinal
SCS	How much do you agree with these statements about learning science?	
	I usually do well in science. (SCS1)[a]	Ordinal
	Science is harder for me than for many of my classmates. (SCS2)	Ordinal
	Science is not one of my strengths. (SCS3)	Ordinal
	I learn things quickly in science. (SCS4)[a]	Ordinal
PATS	How much do you agree with these statements about learning science?	
	I enjoy learning science. (PATS1)[a]	Ordinal
	Science is boring. (PATS2)	Ordinal
	I like science. (PATS3)[a]	Ordinal
SA	Five plausible values for each student.	Continuous

PTSR perceived teacher–student relationship, *PPR* perceived peer relationship, *SCS* self-confidence in learning science, *PATS* positive attitude toward science

[a]For these items, student responses were coded in reverse

pairs of the 14 booklets (Foy et al., 2008) and to provide SA scores. Through the multiple imputation method of IRT (Mislevy, 1991), five PVs were obtained for each student. The average of the five PVs indicates the SA expectation value, and the standard deviation of the five PVs refers to the standard error of measurement. In the current study, each participant's five PVs were used to estimate the path coefficients and their measurement errors caused by the measurement of SA in the SRAM. The average reliability coefficient of the science test for Taiwanese grade 8 students was about 0.86.

Data Analysis

A two-step approach (Anderson & Gerbing, 1988) was adopted to process the model testing and fitting, with LISREL version 8.70 (Jöreskog &

Sorbom, 2004). Confirmatory factor analysis (CFA) was used to validate the measurement model, and then the feasible measurement model was used in the SEM to test SRAM.

Ordinal Data. Because the indicators of the four latent constructs (PTSR, PPR, SCS, PATS) belong to ordinal scale and the PVs of SA to interval scale, the asymptotic covariance matrices of all indicators needed to be generated first. Then the weighted least squares estimator, in LISREL referred to as the asymptotically distribution-free estimator, was used to estimate the parameters involved in the measurement model and in SRAM.

Weighting. Weighting is important to the statistical analysis in order to generalize the results of large-scale surveys. The weight of a given case is inversely proportional to the probability of selecting that case from the population. The TIMSS database provides various weightings dependent on the research purpose. Among these, the house weight is recommended for hypothesis testing.

Fit Statistics. Various fit statistics were adopted to evaluate the measurement models and structural models. If a model adequately fits the data, the goodness of fit index should be greater than 0.90, the root mean square error of approximation should be less than 0.08, the comparative fit index should be greater than 0.90, and the non-normed fit index should be greater than 0.90 (Kline, 2010). Because the chi-square test is sensitive to sample size and appropriate only for moderate sample sizes (N = 200 – 500), the chi-square index is improper to test model fitting for samples in large-scale surveys. However, we still provided the critical-N (CN) for reference, which is the estimated sample size required to make the chi-square test significant at the 0.05 level. The model is acceptable if CN is greater than 200 (Hoelter, 1983).

Parameters and Standard Error Estimation. The SA measurement errors are important considerations in structural modeling. Therefore, all the path coefficients were estimated by averaging the coefficients estimated through five replications of the same modeling process; each replication used one of the five sets of PVs as the indicator of SA (Mislevy, 1991). Because TIMSS did not adopt a simple random selection method at the individual level, Foy et al. (2008) recommended a two-stage jackknife replication technique to estimate the standard errors of statistics for a secondary analysis. The current study conducted 80 replications of SEM for each sample group to estimate the path coefficients and their standard errors. Details of the procedure to estimate the coefficients and their standard errors and the rationale of jackknife replication technique are provided in "Appendix 1."

Results and Discussion

The correlation matrices, means, and standard deviations of all the indicators in the SRAM for the three samples (total sample, HAG, and LAG) are provided in "Appendix 2." The results of this study are organized and reported according to procedural issues and the ordered research questions. Discussion is provided for each result as reported.

Procedural Issue 1: Measurement Model

CFA needs to be conducted before the model fitting for the SRAM. The main purpose of CFA is to validate the indicators of the constructs adopted in the current model. Figure 2 presents the CFA results for the total sample. Because only one indicator corresponds to SA, the factor loading is set at 1.00. Other factor loadings between the items and their corresponding constructs are 0.57 – 0.96;

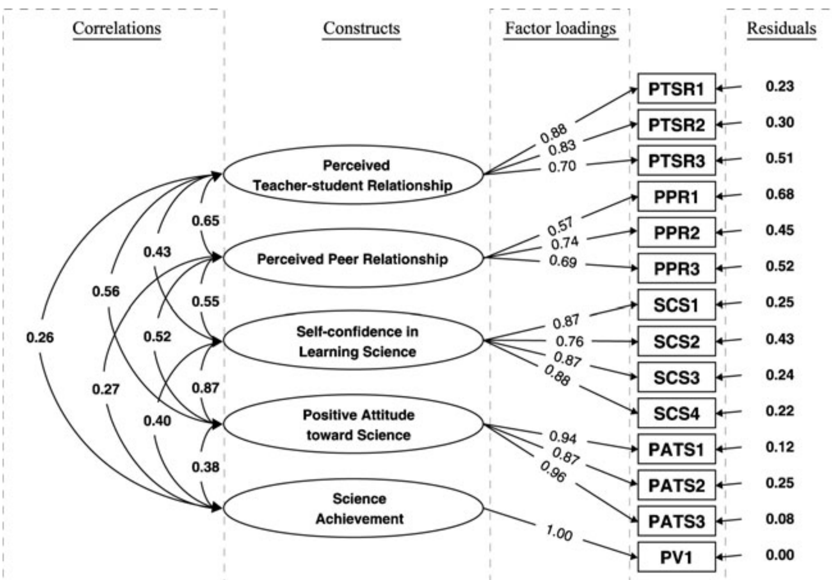

Figure 2. Parameter estimations of confirmatory factor analysis for total sample (n_{ALL} = 3,901)

TABLE 2

Goodness-of-fit indices of the measurement model

Group	*n*	*df*	χ^2	*CN*	*GFI*	*RMSEA*	*CFI*	*NNFI*
High-achieving subgroup	1,956	68	320.17*	606	0.99	0.04	0.99	0.98
Low-achieving subgroup	1,945	68	420.20*	459	0.99	0.05	0.96	0.95
Total sample	3,901	68	624.51*	613	0.99	0.05	0.98	0.97

CN critical number, *GFI* goodness of fit, *RMSEA* root mean square error of approximation, *CFI* comparative fit index, *NNFI* non-normed fit index
*$p < 0.001$

therefore, the overall item quality ranges from good (>0.55) to excellent (>0.71) according to the criteria proposed by Tabachnick & Fidell (2007). The composite reliabilities (ρ_c; Fornell & Larcker, 1981) for PTSR, PPR, SCS, and PATS are 0.94, 0.81, 0.97, and 0.99, respectively, suggesting good internal consistency of items for their construct. CFAs were conducted separately for high- and low-achieving subgroups and showed similar results. Tables 2 and 3 list the fit statistics and composite reliabilities for the total sample and two subgroups. In general, the CFA results confirmed the reliability of the measurement model and its construct validity.

Procedural Issue 2: Structural Model

All the fit indices suggest adequate fit of structural model except for the chi-square tests (Table 4). However, the significance of the chi-square test came from the large sample size effect; it can be remedied by setting the sample size to the CN. These results suggest that SRAM could be used to explain the relationships among the PSR, the

TABLE 3

Composite reliabilities of the four constructs

		Composite reliability			
Group	*n*	*PTSR*	*PPR*	*SCS*	*PATS*
High-achieving subgroup	1,956	0.95	0.80	0.98	0.99
Low-achieving subgroup	1,945	0.92	0.81	0.92	0.98
Total sample	3,901	0.94	0.81	0.97	0.99

PTSR perceived teacher–student relationships, *PPR* perceived peer relationships, *SCS* self-confidence in learning science, *PATS* positive attitude toward science

TABLE 4

Goodness-of-fit indices of the structural model

Group	*n*	*df*	χ^2	*CN*	*GFI*	*RMSEA*	*CFI*	*NNFI*
High-achieving subgroup	1,956	70	334.78*	452.78	0.99	0.04	0.99	0.98
Low-achieving subgroup	1,945	70	436.80*	593.45	0.99	0.05	0.96	0.95
Total sample	3,901	70	648.30*	605.14	0.99	0.05	0.98	0.97

CN critical number, *GFI* goodness of fit, *RMSEA* root mean square error of approximation, *CFI* comparative fit index, *NNFI* non-normed fit index

*$p < 0.001$

affective learning outcomes in science learning, and SA for the total sample and two subgroups.

The path coefficients and their corresponding standard errors for the total sample and subgroups are shown in Figure 3. The path coefficients indicate the effect of the PSR on the affective learning outcomes, the effect of the affective learning outcomes on SA, and the mediating effect of affective learning outcomes between PSR and SA.

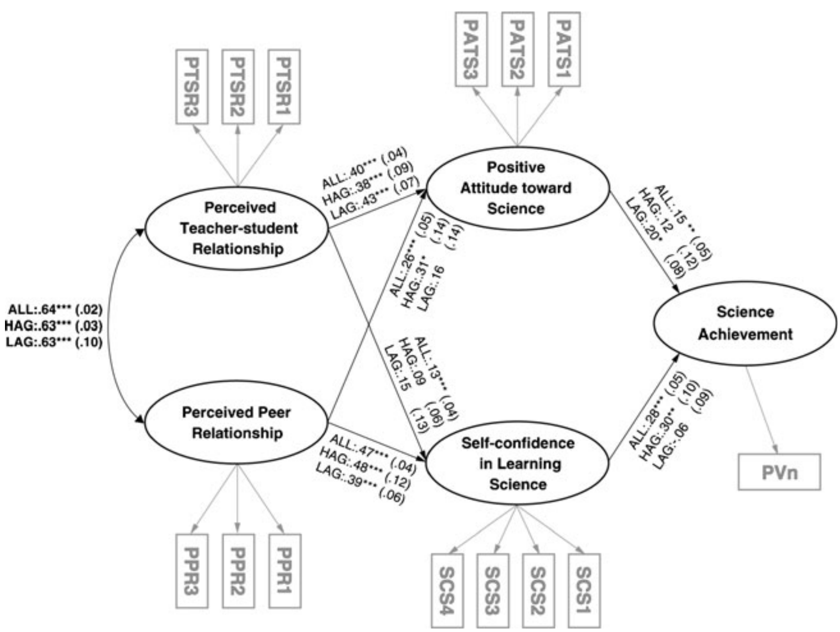

Figure 3. Estimation of standardized path coefficients and standard errors (in *parentheses*) and related items and PVs for total sample group (*ALL*), high-achieving group (*HAG*), and low-achieving group (*LAG*)

Research Question #1: Effects of Perceived Social Relationships on Affective Learning Outcomes

Both types of the perceived social relationships (PTSR and PPR) in science classes for the total sample have positive effects on the students' affective learning outcomes (SCS and PATS). PTSR has a greater effect on PATS ($\gamma_{PTSR \rightarrow PATS} = 0.40$, $p < 0.001$) than on SCS ($\gamma_{PTSR \rightarrow SCS} = 0.13$, $p < 0.01$). Conversely, PPR has a greater effect on SCS ($\gamma_{PPR \rightarrow SCS} = 0.47$, $p < 0.001$) than on PATS ($\gamma_{PPR \rightarrow PATS} = 0.26$, $p < 0.001$). The combination of PTSR and PPR accounts for 31.2 % of the variance in SCS and 36.1 % of the variance in PATS. Therefore, PSR accounts for a considerable portion of the variances in students' affective learning outcomes, consistent with other studies (e.g. Hardre et al., 2006; Nelson & DeBacker, 2008).

PSR effects on the affective learning outcomes are further examined separately for low- and high-achieving subgroups. PTSR for HAG exerts a positive effect on PATS ($\gamma_{PTSR \rightarrow PATS} = 0.38$, $p < 0.001$) and no effect on SCS ($\gamma_{PTSR \rightarrow SCS} = 0.09$, $p > 0.05$) while PPR has positive effects on both SCS ($\gamma_{PPR \rightarrow SCS} = 0.48$, $p < 0.001$) and PATS ($\gamma_{PTSR \rightarrow PATS} = 0.31$, $p < 0.05$). PTSR and PPR together account for 38.9 % of the variance in PATS and 29.3 % of the variance in SCS for the HAG. PTSR for LAG exerts a positive effect on PATS ($\gamma_{PTSR \rightarrow PATS} = 0.43$, $p < 0.001$) whereas no effect on SCS ($\gamma_{PTSR \rightarrow SCS} = 0.15$, $p > 0.05$). Again, PPR exerts a positive effect on SCS ($\gamma_{PPR \rightarrow SCS} = 0.39$, $p < 0.001$) but not on PATS ($\gamma_{PTSR \rightarrow PATS} = 0.16$, $p > 0.05$). PTSR and PPR together account for 24.8 % of the variance in SCS and 29.7 % of the variance in PATS for the LAG. Taken as a whole, PTSR rather than PPR exerts a greater effect on students' PATS, and PPR rather than PTSR has a greater effect on students' SCS across the total sample and the two achieving subgroups.

Research Question #2: Effects of Affective Learning Outcomes on Cognitive Learning Outcome

The results of modeling the total sample's data revealed a medium effect of SCS on SA ($\beta_{SCS \rightarrow SA} = 0.28$, $p < 0.001$) and a small effect of PATS on SA ($\beta_{PATS \rightarrow SA} = 0.15$, $p < 0.01$). SCS and PATS together account for 17.0 % of the variance in SA. Similarly, Chien, Jen & Chang (2008) observed that students' science self-concept predicts their TIMSS SA ($\gamma = 0.24$, $p < 0.001$). Marsh (1990) also found that Australian senior high school students' academic achievement was predicted by their self-concept measured in the previous year ($\gamma = 0.20 - 0.22$,

$p < 0.001$). In Armitage & Conner's (2001) study, attitude predicted an individual's desires and in turn influenced behavior, such as engagement in learning activities. This is consistent with the small direct effect of attitude on academic achievement in the current findings.

The two affective learning outcomes influenced SA of the high- and low-achieving subgroups differently. SCS was a better predictor of SA for HAG ($\beta_{\text{SCS}\rightarrow\text{SA}} = 0.30$, $p < 0.001$) than for LAG ($\beta_{\text{SCS}\rightarrow\text{SA}} = 0.06$, $p > 0.05$). A possible explanation is that low-achieving students' self-evaluation ability was less well developed and less reflective of actual performance than high-achieving students. Low-achieving students often overestimated their academic abilities (Langendyk, 2006). In contrast, PATS is a better predictor of SA for low-achieving students ($\beta_{\text{PATS}\rightarrow\text{SA}} = 0.20$, $p < 0.05$) than for high-achieving students ($\beta_{\text{PATS}\rightarrow\text{SA}} = 0.12$, $p > 0.05$). Ma & Xu (2004) observed that for high-achieving students, the effect of attitude on achievement is not significant. These high performers may engage in learning activities because of other motivations, such as the preparation for the senior high school examination.

Research Question #3: Mediating Effects of Affective Learning Outcomes in SRAM

Table 5 presents the effects of PSR on SA when the two affective learning outcomes (SCS and PATS) are taken as mediators. These outcomes for the total sample significantly mediate the effects of the perceived social relationships (PTSR and PPR) on SA (for PTSR→SCS→SA, $\gamma_{\text{PTSR}\rightarrow\text{SCS}}\beta_{\text{SCS}\rightarrow\text{SA}} = 0.04$, $Z = 2.78$, $p < 0.01$; for PTSR→PATS→SA, $\gamma_{\text{PTSR}\rightarrow\text{PATS}}\beta_{\text{PATS}\rightarrow\text{SA}} = 0.06$, $Z = 2.79$, $p < 0.01$; for PPR→SCS→SA, $\gamma_{\text{PPR}\rightarrow\text{SCS}}\beta_{\text{SCS}\rightarrow\text{SA}} = 0.13$, $Z = 5.07$, $p < 0.001$; for PPR→PATS→SA, $\gamma_{\text{PPR}\rightarrow\text{PATS}}\beta_{\text{PATS}\rightarrow\text{SA}} = 0.04$, $Z = 2.57$, $p < 0.01$). The two PSR accounted for 5.8 % of the variance in SA. When a full model assuming two direct effects of PTSR and PPR on SA was used to fit the data, PTSR and PPR together accounted for 8.3 % of the variance in SA (the same percentage of accountability can be obtained by using the correlations provided in the measurement model in Figure 2). Therefore, 70 % of the total combined effect of PTSR and PPR on SA was explained by the mediating effects, suggesting that SCS and PATS are two important mediators.

Similar analysis and results for the high-achieving subgroup revealed that only SCS significantly mediated the effect of PPR on SA (for

TABLE 5

Mediating effects of affective learning outcome between perceived social relationships and science achievement (standard errors)

			Mediating effects on science achievement[a]	
Group	*n*	*Independent variable*	*Through SCS*	*Through PATS*
High-achieving subgroup	1,956	PTSR	0.03 (0.02)	0.05 (0.05)
		PPR	0.14* (0.06)	0.04 (0.04)
Low-achieving subgroup	1,945	PTSR	−0.01 (0.02)	0.09* (0.04)
		PPR	−0.02 (0.04)	0.03 (0.03)
Total sample	3,901	PTSR	0.04** (0.01)	0.06** (0.02)
		PPR	0.13*** (0.03)	0.04** (0.02)

PTSR perceived teacher–student relationship, *PPR* perceived peer relationship, *SCS* self-confidence in learning science, *PATS* positive attitude toward science

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$

[a]Tests of mediating effect are based on the Aroian test equation

PPR→SCS→SA, $\gamma_{\mathrm{PPR}\rightarrow\mathrm{SCS}}\ \beta_{\mathrm{SCS}\rightarrow\mathrm{SA}} = 0.143$, $Z = 2.39$, $p < 0.05$). However, for the low-achieving subgroup, only PATS significantly mediated the effect of PTSR on SA (for PTSR→PATS→SA, $\gamma_{\mathrm{PTSR}\rightarrow\mathrm{PATS}}\ \beta_{\mathrm{PATS}\rightarrow\mathrm{SA}} = 0.086$, $Z = 2.32$, $p < 0.05$). These results imply that for high-achieving students, peer relationship in science class influences their science achievement through their self-confidence in learning science, whereas for low-achieving students, positive teacher–student relationship improves science achievement through enhancing their attitude toward (learning) science.

Conclusion and Implications

Based on self-determination theory and the emphasis on social relationships in collectivistic cultures (e.g. Asian cultures), a social-relation-based affection-driven model was proposed to examine and explain how Taiwanese grade 8 students at low- or high-achieving levels were influenced by their perceived social relationships in science class through affective learning outcomes. Given the results of previous studies, the items that could be obtained from the TIMSS 2007 data set and its concurrent national survey were accessed according to the necessary constructs, and the directions among these constructs were set a priori before conducting a structure equation modeling.

SEM results confirmed the SRAM in that various perceived social relationships in class influenced science achievement differently, mediated by different affective learning outcomes. SEM on the total sample showed that both the perceived teacher–student and peer relationships predicted students' self-confidence in learning science and their positive attitude toward science as well. However, the PTSR had a greater effect on PATS than on SCS whereas the PPR influenced SCS more than PATS. In addition, both SCS and PATS exerted effects on SA, with SCS a better predictor of SA. Finally, the two affective learning outcomes (SCS and PATS) mediated the effects of the two social relationships (PTSR and PPR) on SA. These SEM results on the total sample suggest that, as being combined SRAM with self-determination theory (Ryan & Deci, 2000b), the fulfillment of the need of relatedness in science class, such as the perception of teacher support and cohesive peer relationships, improved students' self-confidence in learning science and caused a more positive attitude toward science. Therefore, for students from collectivistic cultures, the need of relatedness is so important that the needs of competence and autonomy are influenced, as are their academic performances subsequently.

Another important finding of the current study comes from the mechanisms of learning motivation between high and low achievers. For high achievers, a good PPR exerts an effect on SA through enhancing their self-confidence, whereas for low-achievers, a good PTSR influences SA through having a PATS. PATS also explains more variance in SA than SCS does for low-achievers. These results appear to suggest that low achievers rely more directly on established relationships with their teachers to enhance their learning while high achievers are more likely to have equally high-achieving peers relying on them to enhance their learning.

Educational Implications

Historically, students' self-confidence and attitude toward science had been underemphasized in the secondary school science curriculum in Taiwan. Recently, government institutes have begun to promote some research projects and innovative curricula based on constructivist-oriented science learning and teaching approaches. These educational reforms encourage student-directed inquiries and social interactions that require confidence and collaboration in co-constructing understanding. Therefore, our findings about roles of perceived social relationships in science class in improving students' affective learning outcomes, how to create more friendly learning environments so that students can receive both affective and academic supports from their teachers and have cohesive relationships with their peers, is an important consideration for education policy and decision makers.

The results of comparisons between high- and low-achieving groups imply that a good relationship among peers is important to establish high-achieving students' self-confidence in learning science and, thus, indirectly enhance their science achievement. On the other hand, for low-achieving students, building a good relationship with their science teacher is a good way to enhance a positive attitude toward science, which in turn leads to gradual improvement in science achievement and probably more reliance on positive peer relationships. Therefore, science teachers should use different classroom management strategies to improve the high and low achievers' motivation in learning science.

Methodological Implications

Another aim of the current study was to demonstrate a standard procedure to estimate the standard errors of path coefficients when dealing with data obtained by a complex sampling design. In "Appendix 3," Table 9 presents the path coefficients and their corresponding standard errors in the SRAM for the total sample and the two subgroups. As compared to the standard errors estimated from simple random selection (i.e. unadjusted standard error), the adjusted standard errors (i.e. adjusted standard error) by considering the measurement errors and the complex sampling design were larger, indicating that both the assumption of simple random selection and the ignorance of measurement errors led to an underestimate of standard error. When the standard error is underestimated, the risk of capitalization on chance (type I error) increases as does the inadequate decision about null hypotheses. Therefore, for large-scale survey studies using complex sampling design such as TIMSS and PISA, the procedure of variance estimation proposed in the current study is recommended when conducting a secondary analysis.

Suggestions for Future Studies

The current study adopted a confirmatory rather than verification approach, in that constructs and the directions among constructs are specified a priori before conducting model fitting. Because cross-sectional data provide correlations among variables in nature, prospective longitudinal studies or microgenetic observations are expected to provide the changes in time sequence so that causality among variables in the SRAM can be further assured. In addition, due to the constraints of the TIMSS

2007 data set and its concurrent national survey, the current study adopted only the construct of the social supports in science class. Future studies might explore the support from family members as the relatedness in autonomous engagement, especially under the consideration of collectivistic cultures.

A path whose sampling error estimated by jackknife replication technique (i.e. the fifth column in Table 9 of "Appendix 3") that differs considerably from the error based on the assumption of simple random selection (i.e. the fourth column in Table 9 of "Appendix 3") implies an effect of factors at school level or class level on this path. For example, the magnitudes of the standard error adjustments for the total sample group after jackknife replications for the effects of perceived social relationships on affective learning outcomes ($\gamma_{\mathrm{PTSR}\rightarrow\mathrm{PATS}}$, $\gamma_{\mathrm{PTSR}\rightarrow\mathrm{SCS}}$, $\gamma_{\mathrm{PPR}\rightarrow\mathrm{PATS}}$, $\gamma_{\mathrm{PPR}\rightarrow\mathrm{SCS}}$) were much larger than those of the two affective learning outcomes on science achievement ($\beta_{\mathrm{PATS}\rightarrow\mathrm{SA}}$, $\beta_{\mathrm{SCS}\rightarrow\mathrm{SA}}$). Therefore, the considerable difference indicated that the coefficients were influenced by factors at school level or class level so much that the coefficients would differ if another science class replaced the current class in the stratum of school sampling. When this happens, studies using a multilevel analysis or a hierarchical linear model technique are recommended to explore factor at school level or class level. Similar consideration is applicable for high-achieving and low-achieving subgroups in SRAM (e.g. $\beta_{\mathrm{PATS}\rightarrow\mathrm{SA}}$, $\gamma_{\mathrm{PTSR}\rightarrow\mathrm{PATS}}$, $\gamma_{\mathrm{PPR}\rightarrow\mathrm{SCS}}$, $\gamma_{\mathrm{PPR}\rightarrow\mathrm{PATS}}$ for HAG and $\gamma_{\mathrm{PTSR}\rightarrow\mathrm{SCS}}$, $\gamma_{\mathrm{PPR}\rightarrow\mathrm{SCS}}$ for LAG).

Acknowledgments

This study was made possible through generous grants from the Aim for the Top University project of National Taiwan Normal University funded by the Ministry of Education and the project (NSC97-2511-S-003-045-MY5) funded by National Science Council of Taiwan. The authors also appreciate for the mentoring and editorial assistance from Dr. Larry Yore and Mrs. Shari Yore.

Appendix 1

The path coefficients in SRAM were estimated by averaging the coefficients already estimated through the same modeling process by using different sets of plausible values as the indicator of science achievement (see Eq. 1), and the standard error of each coefficient was the combination of measurement error and sampling error according to the following steps:

$$\widehat{\mu} = \frac{1}{M}\sum_{i=1}^{M}\widehat{\mu}_i \tag{1}$$

Step 1: Estimation of the measurement error

Based on the five sets of coefficients estimated through corresponding sets of students' plausible values, the measurement errors were aggregated according to Eq. 2 (Mislevy, 1991; Foy et al., 2008).

$$\widehat{\sigma}^2_{(\mathrm{PV})} = \frac{1}{M-1}\sum_{i=1}^{M}(\widehat{\mu}_i - \widehat{\mu})^2 \tag{2}$$

In Eqs. 1 and 2, $\widehat{\mu}$ can be any statistic (e.g. mean, correlation, or path coefficients), and M is the number of sets of PVs, which is equal to five here.

Step 2: Estimation of the sampling error

In addition to measurement error, the other source of the variability for path coefficients comes from the sampling error. TIMSS 2007 used a two-stage stratified cluster sampling design. In the first stage, 150 schools were selected according to some variables of interest, such as school type or location. In the second stage, one or two classes in the sampled school were selected at random and all the students in the selected classes were surveyed. Because students in the same class will have the same contextual variables at the class and school levels, the effective sample size could be much less than for the same number of students selected by simple random selection. If we treat the sampled students as though they were sampled through simple random selection, we may underestimate the standard errors of all the coefficients. The two-stage jackknife (JK) replication technique can be utilized to estimate the standard errors caused by the sampling design. In order to conduct the JK replications, theoretically an additional 75 replications should be processed for each set of PVs and the results of 375 replications in total should be aggregated through Eqs. 3 and 4 (Foy et al., 2008).

$$\widehat{\sigma}^2_{(\widehat{\mu}_i)} = \sum_{k=1}^{75}(\mu_{ik} - \widehat{\mu}_i)^2 \tag{3}$$

$$\hat{\sigma}^2_{(\hat{\mu})} = \frac{1}{M}\sum_{i=1}^{M}\hat{\sigma}^2_{(\hat{\mu}_i)} \tag{4}$$

Step 3: Standard error estimation

To estimate the standard errors for all the statistics, the last step is to combine the sampling error and the measurement error portions according to Eq. 5 (Foy et al., 2008).

$$\hat{\sigma}_{(\widehat{\mu\mathrm{PV}})} = \sqrt{\hat{\sigma}^2_{(\hat{\mu})} + \left(1 + \frac{1}{M}\right) \cdot \hat{\sigma}^2_{(\mathrm{PV})}} \tag{5}$$

Due to the fact that the same distribution constraints hold for the five sets of student PVs, in this study only an additional 75 replications for the first set of PVs were conducted in order to estimate the sampling errors for all the coefficients. In other words, $\hat{\sigma}^2_{(\hat{\mu}_1)}$ is utilized instead of $\hat{\sigma}^2_{(\hat{\mu})}$ in Eq. 5.

Appendix 2

TABLE 6

Correlation matrix, means, and standard deviations of the indicators for total sample (n = 3,901)

Item number	*Item code*	*Item number*													
		1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*
1	PTSR1	1.00													
2	PTSR2	0.63	1.00												
3	PTSR3	0.48	0.46	1.00											
4	PPR1	0.27	0.28	0.24	1.00										
5	PPR2	0.33	0.40	0.25	0.38	1.00									
6	PPR3	0.26	0.28	0.29	0.36	0.38	1.00								
7	SCS1	0.28	0.26	0.19	0.18	0.29	0.32	1.00							
8	SCS2	0.13	0.10	0.02	0.06	0.17	0.14	0.43	1.00						
9	SCS3	0.18	0.17	0.06	0.11	0.19	0.21	0.53	0.59	1.00					
10	SCS4	0.28	0.24	0.18	0.18	0.25	0.33	0.67	0.44	0.50	1.00				
11	PATS1	0.40	0.32	0.24	0.16	0.26	0.30	0.62	0.38	0.47	0.66	1.00			
12	PATS2	0.37	0.28	0.21	0.10	0.20	0.20	0.45	0.44	0.50	0.48	0.62	1.00		
13	PATS3	0.42	0.32	0.25	0.16	0.27	0.31	0.62	0.39	0.47	0.67	0.83	0.66	1.00	
14	PV1	0.26	0.20	0.27	0.15	0.18	0.23	0.38	0.20	0.26	0.37	0.37	0.30	0.38	1.00
M		2.68	2.72	3.44	2.75	2.46	2.67	2.50	2.36	2.22	2.25	2.47	2.67	2.47	5.63
SD		0.98	0.85	0.79	0.85	0.91	0.86	0.85	0.94	1.01	0.85	0.93	0.97	0.96	0.88

TABLE 7

Correlation matrix, means, and standard deviations of the indicators for HAG (n_{HAG} = 1,956)

Item number	*Item code*	*Item number* 1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	PTSR1	1.00													
2	PTSR2	0.62	1.00												
3	PTSR3	0.48	0.43	1.00											
4	PPR1	0.25	0.25	0.18	1.00										
5	PPR2	0.28	0.37	0.20	0.35	1.00									
6	PPR3	0.20	0.24	0.18	0.32	0.37	1.00								
7	SCS1	0.20	0.22	0.09	0.17	0.29	0.32	1.00							
8	SCS2	0.18	0.16	0.06	0.13	0.22	0.23	0.60	1.00						
9	SCS3	0.20	0.20	0.07	0.15	0.24	0.26	0.68	0.70	1.00					
10	SCS4	0.21	0.21	0.09	0.16	0.23	0.32	0.66	0.60	0.65	1.00				
11	PATS1	0.36	0.28	0.17	0.15	0.25	0.28	0.58	0.50	0.58	0.63	1.00			
12	PATS2	0.37	0.27	0.19	0.11	0.21	0.20	0.46	0.48	0.53	0.50	0.67	1.00		
13	PATS3	0.38	0.29	0.19	0.16	0.25	0.31	0.57	0.50	0.58	0.64	0.84	0.71	1.00	
14	PV1	0.14	0.15	0.11	0.08	0.15	0.19	0.33	0.29	0.32	0.31	0.29	0.22	0.28	1.00
M		2.88	2.84	3.59	2.84	2.59	2.82	2.79	2.55	2.49	2.54	2.79	2.95	2.79	6.31
SD		0.91	0.79	0.65	0.78	0.87	0.79	0.79	0.93	0.98	0.81	0.87	0.88	0.88	0.46

TABLE 8

Correlation matrix, means, and standard deviations of the indicators for LAG (n_{LAG} = 1,945)

Item number	*Item code*	*Item number* 1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	PTSR1	1.00													
2	PTSR2	0.62	1.00												
3	PTSR3	0.44	0.45	1.00											
4	PPR1	0.27	0.28	0.26	1.00										
5	PPR2	0.33	0.40	0.25	0.39	1.00									
6	PPR3	0.27	0.28	0.33	0.37	0.36	1.00								
7	SCS1	0.25	0.23	0.16	0.15	0.23	0.25	1.00							
8	SCS2	0.03	0.00	-0.07	-0.03	0.07	0.01	0.20	1.00						
9	SCS3	0.07	0.08	-0.03	0.03	0.08	0.10	0.30	0.44	1.00					
10	SCS4	0.24	0.21	0.15	0.15	0.21	0.26	0.60	0.21	0.25	1.00				
11	PATS1	0.36	0.31	0.20	0.11	0.21	0.24	0.57	0.19	0.28	0.60	1.00			
12	PATS2	0.30	0.24	0.15	0.05	0.12	0.13	0.33	0.35	0.40	0.36	0.51	1.00		
13	PATS3	0.38	0.30	0.21	0.10	0.22	0.25	0.57	0.21	0.27	0.62	0.78	0.55	1.00	
14	PV1	0.19	0.12	0.23	0.13	0.09	0.14	0.10	-0.05	-0.05	0.12	0.13	0.08	0.15	1.00
M		2.48	2.59	3.29	2.66	2.32	2.53	2.20	2.18	1.96	1.97	2.17	2.39	2.15	4.96
SD		1.00	0.89	0.88	0.91	0.93	0.90	0.80	0.92	0.97	0.79	0.90	0.98	0.92	0.67

Appendix 3

TABLE 9

Path coefficients and their error estimations of SRAM for the three sample groups

Group	*Estimates*	*Coefficient estimation*	*Unadjusted standard error*[a]	*Sampling error*[b]	*Measurement error*[c]	*Adjusted standard error*[d]
HAG	$\beta_{\text{SCS}\rightarrow\text{SA}}$	0.299**	0.061	0.091	0.036	0.099
	$\beta_{\text{PATS}\rightarrow\text{SA}}$	0.124	0.055	0.115	0.035	0.122
	$\gamma_{\text{PTSR}\rightarrow\text{SCS}}$	0.088	0.040	0.060	0.011	0.061
	$\gamma_{\text{PTSR}\rightarrow\text{PATS}}$	0.382***	0.037	0.093	0.011	0.094
	$\gamma_{\text{PPR}\rightarrow\text{SCS}}$	0.479***	0.043	0.116	0.004	0.116
	$\gamma_{\text{PPR}\rightarrow\text{PATS}}$	0.306*	0.041	0.144	0.004	0.144
	$\phi_{\text{PTSR}\longleftrightarrow\text{PPR}}$	0.628***	0.026	0.033	0.005	0.034
LAG	$\beta_{\text{SCS}\rightarrow\text{SA}}$	-0.059	0.055	0.090	0.019	0.093
	$\beta_{\text{PATS}\rightarrow\text{SA}}$	0.202*	0.057	0.077	0.018	0.079
	$\gamma_{\text{PTSR}\rightarrow\text{SCS}}$	0.150	0.038	0.126	0.005	0.126
	$\gamma_{\text{PTSR}\rightarrow\text{PATS}}$	0.425***	0.037	0.072	0.005	0.072
	$\gamma_{\text{PPR}\rightarrow\text{SCS}}$	0.389***	0.041	0.062	0.002	0.063
	$\gamma_{\text{PPR}\rightarrow\text{PATS}}$	0.157	0.041	0.144	0.003	0.144
	$\phi_{\text{PTSR}\longleftrightarrow\text{PPR}}$	0.626***	0.025	0.095	0.002	0.095
ALL	$\beta_{\text{SCS}\rightarrow\text{SA}}$	0.276***	0.048	0.044	0.019	0.049
	$\beta_{\text{PATS}\rightarrow\text{SA}}$	0.146**	0.049	0.048	0.018	0.052
	$\gamma_{\text{PTSR}\rightarrow\text{SCS}}$	0.125**	0.028	0.040	0.002	0.040
	$\gamma_{\text{PTSR}\rightarrow\text{PATS}}$	0.401***	0.026	0.041	0.002	0.041
	$\gamma_{\text{PPR}\rightarrow\text{SCS}}$	0.469***	0.029	0.044	0.002	0.044
	$\gamma_{\text{PPR}\rightarrow\text{PATS}}$	0.259***	0.029	0.046	0.002	0.046
	$\phi_{\text{PTSR}\longleftrightarrow\text{PPR}}$	0.643***	0.017	0.021	0.001	0.021

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$

[a]Standard errors estimated by treating the examinees were sampled through a simple random selection method

[b]Sampling errors estimated by the two-stage jackknife replication method

[c]Measurement errors estimated based on the five sets of plausible values

[d]Standard errors estimation by taking the complex sampling design and the measurement error of science achievement into consideration

References

Anderson, J. & Gerbing, D. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103*, 411–423.

Armitage, C. J. & Conner, M. (2001). Efficacy of the theory of planned behaviour: A meta-analytic review. *British Journal of Social Psychology, 40*, 471–499.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191–215.

Bao, X. H. & Lam, S. F. (2008). Who makes the choice? Rethinking the role of autonomy and relatedness in Chinese children's motivation. *Child Development, 79*(2), 269–283.

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

Chien, C.-L., Jen, T.-H. & Chang, S.-T. (2008). Academic self-concept and achievement within and between math and science: An examination on Marsh and Köller's unification model. *Bulletin of Educational Psychology, 40*, 107–126.

Cosmovici, E. M., Idsoe, T., Bru, E. & Munthe, E. (2009). Perceptions of learning environment and on-task orientation among students reporting different achievement levels: A study conducted among Norwegian secondary school students. *Scandinavian Journal of Educational Research, 53*(4), 379–396.

Dorman, J. P. (2001). Associations between classroom environment and academic efficacy. *Learning Environments Research, 4*, 243–257.

Erlauer, L. (2003). *The brain-compatible classroom: Using what we know about learning to improve teaching*. Alexandria, VA: Association for Supervision and Curriculum Development.

Fornell, C. & Larcker, D. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*, 39–50.

Foy, P., Galia, J. & Li, I. (2008). Scaling the data from the TIMSS 2007 mathematics and science assessments. In J. F. Olson, M. O. Martin & I. V. S. Mullis (Eds.), *TIMSS 2007 technical report* (pp. 225–279). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Hardre, P. L., Chen, C.-H., Huang, S.-H., Chiang, C.-T., Jen, F.-L. & Warden, L. (2006). Factors affecting high school students' academic motivation in Taiwan. *Asia Pacific Journal of Education, 26*, 189–207.

Hoelter, J. W. (1983). The analysis of covariance structures: Goodness-of-fit indices. *Sociological Methods Research, 11*, 325–344.

Jöreskog, K. G. & Sorbom, D. (2004). *LISREL 8: User's reference guide*. Lincolnwood, IL: Scientific Software International.

Keller, H. (2012). Autonomy and relatedness revisited: Cultural manifestations of universal human needs. *Child Development Perspectives, 6*(1), 12–18.

Kline, R. B. (2010). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford.

Langendyk, V. (2006). Not knowing that they do not know: Self-assessment accuracy of third-year medical students. *Medical Education, 40*, 173–179.

Ma, X. & Xu, J. (2004). Determining the causal ordering between attitude toward mathematics and achievement in mathematics. *American Journal of Education, 110*, 256–280.

Markus, H. R. & Kitayama, S. (2003). Culture, self, and the reality of the social. *Psychological Inquiry, 14*(3), 277–283.

Marsh, H. W. (1990). Causal ordering of academic self-concept and academic achievement: A multiwave, longitudinal panel analysis. *Journal of Educational Psychology, 82*, 646–656.

Martin, M. O. & Preuschoff, C. (2008). Creating the TIMSS 2007 background Indices. In J. F. Olson, M. O. Martin & I. V. S. Mullis (Eds.), *TIMSS 2007 technical report* (pp. 225–279). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

McCoach, D. B., Black, A. C. & O'Connell, A. A. (2007). Errors of inference in structural equation modeling. *Psychology in the Schools, 44*(5), 461–470.

Mislevy, R. J. (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika, 56*(2), 177–196.

Multon, K. D., Brown, S. D. & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology, 38*, 30–38.

Nelson, R. M. & DeBacker, T. K. (2008). Achievement motivation in adolescents: The role of peer climate and best friends. *The Journal of Experimental Education, 76*, 170–189.

Ng, F. F. Y., Kenney-Benson, G. A. & Pomerantz, E. M. (2004). Children's achievement moderates the effects of mothers' use of control and autonomy support. *Child Development, 75*(3), 764–780.

Ryan, R. M. & Deci, E. L. (2000a). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Education Psychology, 25*, 54–67.

Ryan, R. M. & Deci, E. L. (2000b). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*, 68–78.

Shrigley, R. L. (1990). Attitude and behavior are correlates. *Journal of Research in Science Teaching, 27*, 97–113.

Tabachnick, B. G. & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn & Bacon.

Taylor, I. M. & Lonsdale, C. (2010). Cultural differences in the relationships among autonomy support, psychological need satisfaction, subjective vitality, and effort in British and Chinese physical education. *Journal of Sport & Exercise Psychology, 32*(5), 655–673.

Tsai, C.-C. & Kuo, P. C. (2008). Cram school students' conceptions of learning and learning science in Taiwan. *International Journal of Science Education, 30*, 353–375.

United States National Research Council (2007). *Taking science to school: Learning and teaching science in grades K–8*. Washington, DC: National Academies Press.

Weinburgh, M. (1995). Gender differences in student attitudes toward science: A meta-analysis of the literature from 1970 to 1991. *Journal of Research in Science Teaching, 32*(4), 387–398.

Wolf, S. J. & Fraser, B. J. (2008). Learning environment, attitudes and achievement among middle-school science students using inquiry-based laboratory activities. *Research in Science Education, 38*, 321–341.

Yore, L. D., Anderson, J. O. & Chui, M.-H. (2010). Moving PISA results into the policy arena: Perspectives on knowledge transfer for future considerations and preparations. *International Journal of Science and Mathematics Education, 8*(3), 593–609.

Tsung-Hau Jen, Che-Di Lee and Kuan-Ming Chen
Science Education Center
National Taiwan Normal University
No. 88, Sec. 4th, Ting-chou Rd., Taipei 11677, Taiwan, Republic of China
E-mail: tsunghau@ntnu.edu.tw

Che-Di Lee
E-mail: chedi.lee@ntnu.edu.tw

Kuan-Ming Chen
E-mail: kmchen@ntu.edu.tw

Chin-Lung Chien
Department of Psychology
National Chengchi University
Taipei, Taiwan
E-mail: 95752501@nccu.edu.tw

Ying-Shao Hsu
Graduate Institute of Science Education
National Taiwan Normal University
Taipei, Taiwan
E-mail: yshsu@ntnu.edu.tw