

Running head: SCHOOL ATTITUDE ASSESSMENT

Psychometric Properties of the School Attitude Assessment Survey-Revised with International
Baccalaureate (IB) High School Students

Robert F. Dedrick, Elizabeth Shaunessy-Dedrick, Shannon M. Suldo, and John Ferron
University of South Florida

This is a pre-copyedited, author-produced PDF of an article accepted for publication in Gifted Child Quarterly following peer review. The version of record [citation information below] is available online at <https://journals.sagepub.com/doi/10.1177/0016986214559596>

Dedrick, R., Shaunessy-Dedrick, E., Suldo, S. M., & Ferron, J. (2015). Psychometric properties of the School Attitude Assessment Survey- Revised with International Baccalaureate high school students. *Gifted Child Quarterly*, 59(1), 38 – 54. doi: 10.1177/0016986214559596

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A100911 to The University of South Florida. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Correspondence concerning this article should be addressed to Robert F. Dedrick, Ph.D., Department of Educational and Psychological Studies, University of South Florida, 4202 East Fowler Avenue, EDU 105, Tampa, FL, 33620. Phone: (813) 974-5722. E-mail: dedrick@usf.edu

Abstract

In two studies ($n_s = 312$ and 1149) with 9–12 grade students in pre-International Baccalaureate (IB) and IB Diploma programs, we evaluated the reliability, factor structure, measurement invariance, and criterion-related validity of the scores from the School Attitude Assessment Survey-Revised (SAAS-R; McCoach & Siegle, 2003a). Reliabilities of the five SAAS-R subscale scores were good ($\alpha_s > .80$) for pre-IB (grades 9-10) and IB students (grades 11-12). Study 1 model fit indices for the five-factor SAAS-R model from confirmatory factor analyses showed greater misfit than those previously reported by McCoach and Siegle. In contrast, Study 2 fit indices for the five-factor model with pre-IB and IB students were similar to values reported by McCoach and Siegle. Tests of measurement invariance in Study 2 using multigroup confirmatory factor analysis identified three items within the Motivation/Self-Regulation subscale that differed in their item intercepts (i.e., uniform differential item functioning) with pre-IB students endorsing these items more strongly compared to IB students. Based on these results along with evidence of criterion-related validity as reflected in the moderate statistical relations between the SAAS-R subscales and students' GPAs, the SAAS-R shows promise as a research tool that can be used to examine the psychological factors associated with pre-IB and IB students' academic achievement.

Keywords: International Baccalaureate, confirmatory factory analysis

Psychometric Properties of the School Attitude Assessment Survey-Revised with International
Baccalaureate (IB) High School Students

The School Attitude Assessment Survey-Revised (SAAS-R; McCoach & Siegle, 2003a), a revision of McCoach's School Attitude Assessment Survey (2002), is a 35-item survey instrument designed to measure five psychological factors associated with students' academic achievement. The five factors are Academic Self-Perceptions, Attitudes Toward Teachers, Attitudes Toward School, Goal Valuation, and Motivation/Self-Regulation. The SAAS-R, which uses a seven-point response scale (1 = *strongly disagree* to 7 = *strongly agree*), has been used with diverse groups of students, including academically gifted achievers and underachievers (e.g., Matthews & McBee, 2007), with the five factors serving various roles in research studies (e.g., outcome variable, predictor, mediator). A summary of studies that have used the SAAS-R is available online as supplemental materials (Appendix A).

McCoach and Siegle (2003a), the developers of the measure, provided initial psychometric support of the SAAS-R using confirmatory factor analysis (CFA) and comparative analyses that tested differences on the means of the five SAAS-R factors between gifted high achievers and gifted underachievers. Suldo, Shaffer, and Shaunessy (2008) added to the psychometric support for the SAAS-R using exploratory and confirmatory factor analyses, correlational analyses of theoretically related variables (e.g., school climate, school satisfaction, academic self-efficacy, in-school conduct, time spent on homework), and comparative analyses that tested differences on the means of the five SAAS-R factors among high school students with low, average, and high achievement. Recently, Davie (2012) used CFA and provided support for the five-factor model underlying the SAAS-R, finding acceptable model fit (i.e., comparative fit index [CFI] = .94, Tucker-Lewis Index [TLI] = .94, root mean square error of approximation [RMSEA] = .05, and

standardized root mean square residual [SRMR] = .04) in a sample of 847 high school students (gifted high achievers, gifted underachievers, and non-gifted low achievers).

Based on these initial results, the SAAS-R has shown promise as a research tool for evaluating the psychological factors associated with students' academic achievement. This tool is much needed as researchers aim to better understand the experiences of various student populations including high-achieving high school students in specialized programs such as the International Baccalaureate (IB) Diploma program. The IB program, developed in the late 1960s, is designed for high school juniors and seniors but schools often offer students in grades 9 and 10 a structured pre-IB curriculum (International Baccalaureate Organization [IBO], 2013a). The IB program is unique in its holistic approach to education and its emphasis on the well-rounded, civically-minded, and globally-conscious learner (IBO, 2013e). Like Advanced Placement classes and dual enrollment courses, the IB Diploma program is one of the accelerated programming options offered to gifted learners in secondary settings (Colangelo, Assouline, & Gross, 2004), though the IB program was not initially designed for this population exclusively. While there is a paucity of research documenting the effectiveness of the IB program in meeting gifted learners' cognitive and affective needs, Hertberg-Davis and Callahan (2008) reported that many IB students felt both challenged and overwhelmed by the rigorous coursework.

The IB Diploma program has grown worldwide and in the United States. The Americas region of IB, including North, Central, and South America, offered approximately 800 programs in 2000 and more than 2300 in 2012 (IBOb). In 2013, 800 IB Diploma programs were offered in the United States (IBO, 2013d); the largest concentration of IB Diploma programs is in the Southeastern states, with the most IB Diploma programs offered in Florida (Perna et al., 2011). Many colleges recognize the IB Diploma and award college credits accordingly, while others

award additional weight in college application decisions for IB diploma completers or by end-of-course exam scores. For instance, in Florida, students who graduate with an IB Diploma may begin college at public universities with sophomore standing (IBO, 2013c).

As researchers begin extending the use of the SAAS-R to various subgroups of high-achieving students, such as those in IB programs, it is critical that the psychometric properties of the scores from the SAAS-R are evaluated and reported when the instrument is used in different contexts with different populations. One important psychometric property of an instrument is its factor structure, which represents the theoretical dimensions underlying the instrument and the grouping of items that reflect these dimensions or factors. Evaluation of an instrument's factor structure has played a central role in the construct validation process (DiStefano & Hess, 2005). Reliability (e.g., Cronbach's alpha) is another important psychometric property that needs to be assessed. Finally, for the SAAS-R, evidence of criterion-related validity is particularly relevant because the instrument was designed to differentiate "academically able achievers from academically able underachievers" (McCoach & Siegle, 2003a, p. 425).

The American Educational Research Association (AERA), American Psychological Association (APA), and the National Council on Measurement in Education (NCME) collectively assert in the *Standards for Educational and Psychological Testing* (2014) that it is inappropriate to assume that measurement properties such as the factor structure, reliability, and criterion-related validity that were identified during the development of an instrument will necessarily generalize to other populations and contexts, and as a result call for a reexamination of measurement quality when an instrument is used in a new context or with different groups. This call is particularly important given that studies that have examined the psychometric properties of various psychological and educational measures have often found differences in

score reliability (Vacha-Haase & Thompson, 2011) and factor structures (Sass, 2011) when measures are administered under different conditions, with these differences potentially compromising the researchers' statistical analyses (e.g., multiple regression, ANOVA) in terms of attenuated relationships and biased parameter estimates.

Despite calls for more extensive reporting of psychometric information (e.g., see Kieffer, Reese, & Vacha-Haase's [2010] call in the context of giftedness research), many researchers have not examined the measurement properties of instruments when used with their research samples and instead have relied on prior results reported during the original development of the measure (Thompson & Vacha-Haase, 2000). In the 38 studies that have used the SAAS-R (see supplemental materials), only two (Davie, 2012; Suldo et al., 2008) examined the factor structure of the measure, and only Davie (2012) conducted a confirmatory factor analysis (CFA) at the item-level (Davie's sample was 847 high school students at one school). Suldo et al. used exploratory factor analysis to examine the 35-item SAAS-R at the item level and then conducted a CFA of 15 item parcels created from the 35 items (e.g., from the seven Academic Self-Perceptions items, three composite scores or parcels were created by averaging groups of items consisting of three, two, and two Academic Self-Perceptions items) to evaluate the five-factor model underlying the SAAS-R. Interestingly, of the 38 studies that used the SAAS-R, 24 (63%) calculated Cronbach's alpha reliabilities for the data collected in the study while the remaining studies either relied on the reliabilities reported by McCoach and Siegle (2003a) or did not mention reliability. The limited reliability reporting for the SAAS-R is consistent with Vacha-Haase and Thompson's (2011) review of 12 years of reliability generalization studies, which found that 54.6% of the 12,994 primary studies that were part of their review did not mention reliability and among those that did mention reliability, 15.7% relied on reliability estimates

provided in test manuals or prior articles. Warne, Lazo, Ramos and Ritter (2012) found similar results in their review of statistical methods used in five gifted education research journals between 2006 and 2010. Warne et al. (2012) reported that only 53.3% of the quantitative and mixed methods articles provided reliability estimates for their own data (Cronbach's alpha was the most widely reported reliability estimate). The limited reporting of measurement details goes counter to the guidelines provided by AERA (Duran et al., 2006) for reporting empirical research, which state that "sufficient detail should be provided to make clear that measures are being used appropriately, have suitable dependability (reliability) properties, and are interpreted properly for the groups studied" (p. 36).

Aligned with *AERA's Standards for Reporting on Empirical Social Science Research in AERA Publications* (Duran et al., 2006) and the *Standards for Educational and Psychological Testing* (2014) that call for a reexamination of measurement quality when an instrument is used in a new context or with new samples of students, we used data from two studies to evaluate the factor structure, measurement invariance, reliabilities, and criterion-related validity of the scores from the SAAS-R. We examined the factor structure using confirmatory factor analysis. In addition to examining the factor structure of the SAAS-R separately for the pre-IB and IB groups, we evaluated the equivalence or invariance of the psychometric properties of the SAAS-R (e.g., item intercepts, item factor loadings, item residuals) across the pre-IB and IB groups. Although the issue of measurement invariance has received increased attention outside the field of gifted education (Dimitrov, 2010; Millsap, 2011), Warne et al.'s (2012) review of articles in five gifted journals identified only five studies that conducted tests of invariance (e.g., Peters & Gentry, 2010). We located only one study that examined measurement invariance for the SAAS-R. McCoach and Siegle (2003b) focused on one subscale from the SAAS-R (Academic Self-

Perceptions) and tested the equivalence of the item factor loadings for gifted high school students (grades 9-12) and 9th grade students from a general school population. They found statistically significant differences on two of the seven factor loadings.

Evaluation of measurement invariance is important because if the psychometric properties of items (e.g., item factor loadings, which represent the relation between an item and a factor; item intercepts, which represent the predicted item response when the latent variable is zero) are different across groups, comparisons of mean scores on the latent variables may not be valid. In addition to playing an important part in the validation process, evaluation of measurement invariance may also provide substantive information about how pre-IB and IB groups may interpret and conceptualize the attitudinal items on the SAAS-R. Evidence of very different factor structures for these groups of students may suggest that the meaning of the items is fundamentally different between these groups. Analyses at the individual item level can help identify subtle differences between these groups that might not be seen if analyses focused only on the subscale means of the SAAS-R. These statistical analyses may provide a unique window into the educational experiences of pre-IB and IB students that go beyond simply comparing mean differences on the subscales (e.g. ANOVA).

Lastly, we evaluated the reliability of the five SAAS-R subscale scores using Cronbach's alpha, and for criterion-related validity, we examined the relationships between the five SAAS-R factors and students' grade point averages (GPAs) obtained from school records (details about these analyses are provided in the statistical analysis section). All analyses were conducted in two samples (Studies 1 and 2) of 11th and 12th grade high school students in IB programs and in two samples (Studies 1 and 2) of 9th and 10th graders in a pre-IB curriculum offered at IB Diploma-granting schools.

Method

This research on the SAAS-R is part of a larger project examining stress and coping of high-achieving high school students. The methods and results from two studies of the SAAS-R from the larger project are reported below.

Participants

Study 1. Participants consisted of 312 students (161 pre-IB 9th and 10th graders, and 151 IB 11th and 12th graders) from three public high schools recruited from three school districts in one state in the southeastern part of the United States. Students identified as IB for the current study had to be officially admitted by the school to the IB program (which, in these schools, includes pre-IB students in grades 9 and 10, and IB Diploma students in grades 11 and 12). Consideration for admission to the pre-IB program was based on competitive applications ranked in part according to statewide assessment scores and grade point average in grades seven and eight. Only students who had received written parental permission and had provided signed assent participated in this study. Sample sizes of student participants from the three schools were 85, 102, and 125. The distribution of IB and pre-IB students did not differ by school, $\chi^2(2, N = 312) = 4.29, p = .12$.

The sample was primarily female (61.2%) and White (55.3%). Twelve percent of the students were eligible for Free or Reduced Price Lunch. Seventy-five percent of the fathers and 73.3% of the mothers of the participants had completed college or beyond. There were no statistically significant differences ($p > .05$) between pre-IB and IB students on any of the demographic characteristics in Table 1.

Study 2. Participants consisted of 1149 students (589 pre-IB 9th and 10th graders, and 560 IB 11th and 12th graders) from 10 public high schools recruited from five school districts in one

state in the southeastern United States. Two of the schools that were used in Study 1 participated in Study 2. Different samples of students were used in Study 2 from these two schools (i.e., no students participated in both Studies 1 and 2). As in Study 1, students were considered for IB program admission based on prior achievement. Of the 10 IB sites, eight required a pre-IB curriculum and two other sites (within the same school district) offered the IB Middle Years Program (MYP; IBO, 2013f) for students in grades 9 and 10. Only students who had received written parental permission and had provided signed assent participated. Sample sizes of student participants from the 10 schools ranged from 78 to 169 ($M = 115$). There was a statistically significant difference in the distribution of IB and pre-IB students by school, $\chi^2(9, N = 1149) = 43.05, p < .001$. The largest difference was between one school in which the percentage of participants who were in IB was 64% (36% in pre-IB) and a second school in which 37% were in IB (63% in pre-IB).

Participants were primarily female (59.4%) and White (49.2%). Sixty-five percent of the fathers and 71.0% of the mothers of the participants had completed college or beyond. There were no statistically significant differences ($p > .01$) between pre-IB and IB students on any of the demographic variables in Table 1.

Procedure

Procedures used to recruit schools and student participants were similar across the two studies. Upon obtaining approval from the local school districts and the University Institutional Review Board (IRB), each participating school sent parents consent forms via students served in two classrooms of IB students per grade level; schools, in consultation with cooperating teachers, selected the classes. Four of the five districts' research policies allowed for student incentives for participation (movie ticket passes or \$10 iTunes gift cards). A research team of graduate students

led by two faculty members who served as Principal Investigator (PI) and Co-PI for the project collected data in the spring of 2011 and spring 2012 (for Studies 1 and 2, respectively). All research assistants received training to ensure standardization across data collection. Prior to questionnaire completion, researchers read the student assent form aloud and students completed the assent form before proceeding to the questionnaire. Student participants completed a 14-page questionnaire for Study 1 and a 16-page questionnaire for Study 2, which included demographic items (e.g., gender, grade level), the SAAS-R, and several psychological measures, administered in groups of approximately 10 to 120 (typical was 50-60) during the school day. Completion of the survey packet took approximately 45 minutes. All measures had been piloted with IB students from other schools prior to administration.

Students' unweighted cumulative high school grade point averages (GPA) were obtained directly from the school districts (Study 1) or calculated by the research team based on transcripts from one semester (Spring 2012) obtained from the school districts (Study 2). Grade point averages were recorded on a four-point scale.

School Attitude Assessment Survey-Revised (SAAS-R)

The SAAS-R is a 35-item self-report questionnaire that measures five factors: Academic Self-Perceptions (7 items), Attitudes Toward Teachers (7 items), Attitudes Toward School (5 items), Goal Valuation (6 items), and Motivation/Self-Regulation (10 items). The response scale ranges from 1 (*strongly disagree*) to 7 (*strongly agree*). McCoach and Siegle (2003a) used an iterative process to extend and revise the original School Attitude Assessment Survey (McCoach, 2002). Confirmatory factor analyses of the final 35-item version of the SAAS-R in a combined sample of 645 students (146 11th and 12th graders attending a summer program for talented students, 200 9th grade students in an urban high school, and 299 high school students of varying

achievement levels) indicated acceptable fit of the five-factor model, $\chi^2(550, N = 537) = 1,581.7$ (CFI = .91, TLI = .92, RMSEA = .06, SRMR = .06). Correlations of the five factors ranged from .27 (Academic Self-Perceptions and Attitudes Toward School) to .74 (Goal Valuation and Motivation/Self-Regulation).

For Study 1, we used the original wording on the SAAS-R. For Study 2, we inserted the word *IB* before the words *classes* or *teachers* for the seven items in the Attitudes Towards Teachers subscale (e.g., “My *IB* classes are interesting”; “I relate well to my *IB* teachers”). We made these changes based on feedback from students in Study 1 regarding some ambiguity surrounding whether the term *teacher* referred to those who taught courses in their IB program or other non-IB courses available in their school (e.g., electives, additional college-level courses).

Statistical Analysis

Confirmatory factor analysis (CFA) was used to evaluate the five-factor model underlying the SAAS-R separately for pre-IB and IB students in each study. Confirmatory factor analyses were based on the matrix of Pearson product moment correlations of the 35 SAAS-R items and were conducted in *Mplus* 7.11 (Muthén & Muthén, 1998-2012) using robust maximum likelihood estimation (MLR) in which standard errors and the chi-square test statistic are robust to non-normality (see online supplemental materials, Appendix B, Tables B1 and B2 for measures of skewness and kurtosis for the 35 items of the SAAS-R for Studies 1 and 2, respectively). Full-information maximum likelihood estimation within *Mplus* was used to handle missing data. The amount of missing data was minimal in each study. In Study 1, 94.4% of the pre-IB students had complete data on the 35-item SAAS-R (missing data per item ranged from 0% to <1%), and 95% of the IB students had complete data (missing data per item ranged from 0% to 1.3%). In Study 2, 92% of the pre-IB students had complete data (missing data for the

items ranged from 0 to < 1%); 95% of the IB students had complete data on the SAAS-R with missing data ranging from 0 to 1.1% for the items. One factor loading within each factor (reference indicator) was fixed to one to statistically identify the model.

Fit of the models within each group (i.e., pre-IB and IB) was evaluated using the Satorra-Bentler (SB) scaled (mean-adjusted) chi-square test, Bentler's (1992) normed comparative fit index (CFI), Tucker-Lewis Index (TLI), standardized root mean square residual (SRMR), and the root mean square error of approximation (RMSEA). Hu and Bentler's (1999) cutoff values of greater than or equal to .95 for the CFI and TLI, SRMR less than .08, and RMSEA less than .06 were used as general indicators of acceptable fit of the models. Evaluation of model fit is complex because of the multiple factors that may affect the fit indices (e.g., sample size, model complexity, magnitude of the correlations of the baseline model; see Sun, 2005), and although guidelines have been presented in the literature, many have questioned universal cutoff values and have proposed the use of multiple indices and human judgment that combines theoretical and statistical criteria (Chen, Curran, Bollen, Kirby, & Paxton, 2008). For example, Brown (2006) has suggested that values of the CFI and TLI in the .90 to .95 range may indicate acceptable fit if other fit measures provide evidence of good model fit.

Measurement invariance of the five-factor SAAS-R model between the pre-IB and IB groups was tested within each study using multigroup confirmatory factor analysis (MCFA). MCFA involves testing a series of hierarchically ordered models of increasing restrictiveness. The first model tested was the least restrictive invariance model in which the same five factors underlying the SAAS-R were associated with the same items across the pre-IB and IB groups with no equality constraints imposed on the factor loadings, item intercepts, item unique variances, or factor variances and covariances across groups. This model (Model 1) provides a

test of configural invariance. Metric invariance (Model 2) addresses whether or not the unstandardized factor loadings are the same across the groups. If the factor loadings representing the relationships between the items and the latent variables (e.g., Academic Self-Perceptions) are different across groups, this suggests that the items may have different meanings across groups. This lack of equivalence, indicating the absence of measurement invariance, is often referred to as differential item functioning (DIF), specifically non-uniform DIF (i.e., item response differences between groups vary across the levels of the latent variable). Next, we explored the presence of scalar invariance (Model 3). Scalar invariance was evaluated by examining group differences in the item intercepts (i.e., the extent to which students endorse an item). A lack of equivalence in the intercepts is an indication of uniform DIF (i.e., after equating the groups on a latent variable, one group's item responses differ in the same direction from the other group's responses across all levels of the latent variable). After conducting tests of configural, metric, and scalar invariance, we examined invariance of the item uniqueness (residuals) parameters (Model 4). Finally, we examined structural invariance in terms of the equality of the factor variances (Model 5) and factor covariances (Model 6).

The strategy used to examine the various levels of measurement invariance was to evaluate the Satorra-Bentler (SB) scaled chi-square change ($\Delta SB \chi^2$) for statistical significance relative to the change in degrees of freedom (Δdf) for the models being compared. A scaled chi-square difference is required because "a difference between two SB χ^2 values is not distributed as χ^2 " (Brown, 2006, p. 385). These tests were supplemented by comparing the changes in the CFI, TLI, RMSEA, and SRMR to the guidelines presented by Chen (2007). When the changes in these fit indices for the more restrictive model met Chen's guidelines ($\Delta CFI < .01$, $\Delta TLI < .01$, $\Delta RMSEA < .015$, and $\Delta SRMR < .03$ for the factor loadings, and $\Delta CFI < .01$, $\Delta TLI < .01$,

$\Delta\text{RMSEA} < .015$, and $\Delta\text{SRMR} < .01$ for the item intercepts), we concluded that the hypothesis of invariance was tenable (i.e. do not reject the null hypothesis of equality).

Following the examination of the factor structure (dimensionality) of the SAAS-R, we calculated Cronbach's alphas for each of the SAAS-R subscales by IB group and study. Finally, we used structural equation modeling (SEM) to evaluate criterion-related validity by examining the relationships between the five SAAS-R factors (predictor variables) and the outcome variable of students' grade point averages (GPA) obtained from students' records. As an additional analysis to evaluate the criterion-related validity of the SAAS-R scores, we followed the approach used by McCoach and Siegle (2003a), which involved comparing the observed scores on the five SAAS-R factors for academically able achievers and underachievers.

Results

Descriptive Analyses

Study 1. Based on the descriptive statistics for the 35 SAAS-R items by IB group (pre-IB and IB), the two most strongly endorsed items by the pre-IB group were items 15 ($M = 6.65$, $SD = 0.82$, on a 7-point scale) and 18 ($M = 6.64$, $SD = 0.69$). Both items are within the Goal Valuation factor. For the IB group the same two items were the most strongly endorsed ($M = 6.65$, $SD = 0.74$ and $M = 6.54$, $SD = 0.90$, respectively). Several of the items within the Academic Self-Perceptions and Goal Valuation factors exhibited severe departures from normality (e.g., item 15 within Goals).

Descriptive comparisons of the pre-IB and IB groups on the individual items revealed that the effect sizes (Cohen's d) ranged from -0.32 (item 16 within Attitudes Toward Teachers), where IB students more strongly endorsed this item to an effect of 0.48 (item 23 within Attitudes Toward School), with pre-IB students more strongly endorsing this item. The mean effect size

for the 35 items was 0.01 (median = 0.01) with only 3 of the 35 items having an effect size (absolute value) of 0.35 or greater (item 7, 0.40; item 19, 0.39; and item 23, 0.48). All three of these items were in the Attitudes Toward School subscale and favored pre-IB students; see Table B1 in Appendix B in online supplemental materials).

Study 2. The pattern of descriptive results for Study 2 was similar to that of Study 1. The two most strongly endorsed items by the pre-IB and IB groups were two Goal Valuation items (15 and 18). Effect sizes (Cohen's *d*) ranged from -0.12 (item 5 within Academic Self-Perceptions), where IB students more strongly endorsed this item to 0.34 (item 6 within Attitudes Toward School) with pre-IB students more strongly endorsing this item. The mean effect size for the 35 items was 0.10 (median = 0.11), which was comparable to that in Study 1.

Confirmatory Factor Analysis

Study 1. Model fit indices for the confirmatory factor analyses of the five-factor measurement model are presented in Table 2 with a summary of parameter estimates (standardized loadings, intercorrelations between factors) presented in Tables 3 and 4 (see online supplemental materials, Appendix C, for all standardized and unstandardized parameter estimates for all models). Model fit indices for both the pre-IB and IB groups were generally below the acceptable values, and suggested greater misfit than what McCoach and Siegle (2003a) found with their sample of 645 high school students. Within the current sample, greater model misfit was present for the IB group, $SB\chi^2(550, N = 151) = 1043.84, p < .0001$ (CFI = .83, TLI = .82, RMSEA = .08, SRMR = .09).

All of the standardized pattern coefficients (factor loadings) for the SAAS-R items were statistically significantly different from zero ($p < .05$). Average loadings by subscale for the pre-IB group ranged from .59 for the Motivation/Self-Regulation subscale to .81 for the Attitudes

Toward School subscale. For the IB group the average loadings by subscale ranged from .65 for the Attitudes Toward Teacher subscale to .85 for the Attitudes Toward School subscale.

The correlations between the five factors were all statistically significant ($p < .05$; see Table 4). The largest correlation was between Goals and Motivation/Self-Regulation for the pre-IB (.74) and IB (.66) groups. McCoach and Siegle's largest correlation also was between Goals and Motivation/Self-Regulation ($r = .74$). McCoach and Siegle's correlations between the five factors were generally stronger ($M = .48$, median = .41) compared to those in the IB group ($M = .37$, median = .34) and the pre-IB group ($M = .40$, median = .32).

Modification indices (MIs), which are estimations of how much the chi-square would change if a parameter that has been fixed to zero was freely estimated, were used to examine sources of misfit in the model. Based on the modification indices (MIs), a major source of misfit for both the pre-IB and IB groups involved covariances between the measurement errors for pairs of items. For the pre-IB and IB groups there were 15 and 26 pairs of error covariance terms, respectively, that reflected substantial misfit. In theory, the measurement errors for the items should be random and therefore not covary with each other; however, when there are similarities in item content and language for pairs of items, the item errors may covary. In view of the problems associated with exploratory, post hoc model modification (e.g., capitalization on chance; over fitting models that do not replicate; potential biasing effect on other parameter estimates; Brown, 2006), we only added error covariance terms that were conceptually meaningful: items 2 and 5 (Academic Self-Perceptions); items 1 and 34 (Attitudes Toward Teachers); and items 30 and 33 (Motivation/Self-Regulation). These item pairs were contained within the same SASS-R factor, had very similar wording, and were replicated in Study 2. Even with these three additional terms added to the models, the models did not reach an acceptable

level of fit in either group (a precondition for evaluating measurement invariance) and therefore we did not conduct invariance testing within Study 1 (see Study 2 for invariance testing).

Study 2. Model fit indices for the confirmatory factor analyses of the five-factor measurement model are presented in Table 2 with parameter estimates (standardized loadings, intercorrelations between factors) presented in Tables 3 and 4 (see Appendix C in online supplemental materials for all standardized and unstandardized parameter estimates for all models). Model fit indices for the IB group were equal to or slightly better than those in the pre-IB group. Values for the CFI (.92), RMSEA (.05), and SRMR (.06) for the IB group also were slightly better than those reported by McCoach and Siegle (2003a). The TLI for the IB group (.91) was slightly lower than McCoach and Siegle's value (.92).

All of the standardized pattern coefficients (factor loadings) for the SAAS-R items were significantly different from zero ($p < .05$). Average loadings by subscale for the pre-IB group ranged from .69 for the Motivation/Self-Regulation subscale to .86 for the Attitudes Toward School subscale; for the IB group the average loadings by subscale ranged from .73 for the Academic Self-Perceptions and Motivation/Self-Regulation subscales to .88 for the Attitudes Toward School subscale.

The correlations between the five factors were all statistically significant ($p < .05$; see Table 4). The largest correlation was between Goals and Motivation/Self-Regulation for the pre-IB (.63) and IB (.68) groups. McCoach and Siegle's largest correlation also was for Goals and Motivation/Self-Regulation ($r = .74$). McCoach and Siegle's correlations ($M = .48$, median = .41) were similar to those of the pre-IB group ($M = .48$, median = .47) but slightly stronger than those of the IB group ($M = .42$, median = .35).

We examined the modification indices for the pre-IB and IB models and determined that a major source of misfit involved covariances between errors for three pairs of items (items 2 and 5; items 1 and 34; and items 30 and 33). The wording for these three pairs of items was very similar, a factor that often contributes to error covariances (Brown, 2006). We included these three error covariance terms into the model because these modifications were conceptually meaningful. Inclusion of these error covariances produced a statistically significant improvement in the fit of the models ($\Delta SB \chi^2 = 120.05$, $\Delta df = 3$, $p < .0001$ for pre-IB; $\Delta SB \chi^2 = 73.46$, $\Delta df = 3$, $p < .0001$ for IB; see Table 5 for the fit statistics for the pre-IB and IB groups). Minimal changes in the parameter estimates (standardized loadings, intercorrelations between factors) were observed for the modified models.

After fitting the five-factor model with the three error covariance terms for the pre-IB and IB groups separately, we used multigroup CFA to evaluate measurement invariance (i.e., equality of the item factor loadings, item intercepts, and uniquenesses) and structural invariance (i.e., equality of factor variances and factor covariances). Table 5 summarizes the sequence of invariance tests that were conducted. Overall there was support for configural invariance (equal form, Model 1B), $SB \chi^2 (1094, N = 1149) = 2554.54$, $p < .0001$ (CFI = .93, TLI = .92, RMSEA = .05, SRMR = .05). In the next stage of invariance testing, the results for metric invariance (equal factor loadings) were mixed with the change in chi-square ($\Delta SB \chi^2 = 60.72$, $\Delta df = 30$, $p < .001$) indicating statistically significant differences in the factor loadings between the groups, whereas the changes in the alternative measures of fit indicated that the hypothesis of equal loadings was tenable (e.g., $\Delta CFI = .001$). Based on recommendations by Cheung and Rensvold (1999), we followed up the overall metric invariance test to examine invariance of the factor loadings at the individual factor level using the change in chi-square ($\Delta SB \chi^2$) relative to the change in degrees

of freedom (Δdf) for the more restrictive model. Table 5 summarizes the invariance tests for the factor loadings conducted separately for each of the five SAAS-R factors. There were statistically significant differences for the factor loadings for two of the factors: Attitudes Toward School ($\Delta SB \chi^2 = 20.30$, $\Delta df = 4$, $p < .001$) and Motivation/Self-Regulation ($\Delta SB \chi^2 = 27.52$, $\Delta df = 9$, $p < .01$). Follow-up comparisons for the individual items within these two factors using a Bonferroni correction to adjust the alpha level for the multiple comparisons ($.05/13 = .004$) did not yield any statistically significant differences in the item loadings based on the change in chi-square ($p > .004$). Because no specific item loadings were found to be statistically significantly different across groups, all loadings were constrained to be equal for the remaining invariance tests.

Next, we evaluated the equality of the item intercepts or scalar invariance (i.e., the constraint of equal intercepts was added to the constraint of equal loadings). Similar to metric invariance, the results were mixed. The change in chi-square indicated that there were statistically significant differences in the intercepts between the groups, whereas the changes in the fit indices suggested that the hypothesis of equal intercepts was tenable. Invariance tests using the change in chi-square at the individual factor level revealed statistically significant differences for the item intercepts for two of the factors: Academic Self-Perceptions ($\Delta SB \chi^2 = 13.26$, $\Delta df = 6$, $p < .05$) and Motivation/Self-Regulation ($\Delta SB \chi^2 = 34.71$, $\Delta df = 9$, $p < .0001$; see Table 5). Follow-up comparisons for the individual items within these two factors using a Bonferroni correction ($.05/15 = .003$) identified three items within the Motivation/Self-Regulation factor (items 4, 24, 26; see online supplemental materials, Appendix B, for item content) where the intercepts for the pre-IB group (4.83, 5.92, 5.19) were significantly higher (p

< .003) than the intercepts for the IB group (4.51, 5.70, 4.90). The largest difference was 0.32 on the 7-point response scale and was for item 4, “I check my assignments before I turn them in”.

For the remaining invariance tests, the intercepts for items 4, 24, and 26 were allowed to vary across groups. We tested the equality of the item residual variances (uniquenesses) and found no significant differences between the pre-IB and IB groups ($\Delta SB \chi^2 = 49.80$, $\Delta df = 35$, $p > .05$). A test of the equality of the three error covariances between groups also was not statistically significant ($\Delta SB \chi^2 = 1.64$, $\Delta df = 3$, $p > .05$).

In the last stage of testing we evaluated structural invariance by testing the equality of the variances of the five SAAS-R factors and the covariances of these factors. Changes in the chi-square indicated statistically significant differences in the factor variances ($\Delta SB \chi^2 = 15.56$, $\Delta df = 5$, $p < .01$), whereas the changes in the fit indices suggested that the hypothesis of equal factor variances was tenable. When we conducted individual tests of the factor variances using a Bonferroni correction ($.05/5 = .01$), we found statistically significant differences in the variance of the Attitudes Toward School factor ($\Delta SB \chi^2 = 8.11$, $\Delta df = 1$, $p < .01$) with greater variability for the IB students on this factor. We next tested the equality of the covariances between the factors and found significant differences between groups based on the change in chi-square ($\Delta SB \chi^2 = 26.39$, $\Delta df = 10$, $p < .01$); minimal change in the alternative fit indices was observed suggesting that the hypothesis of equal factor covariances was tenable. When we conducted follow-up comparisons of the 10 factor covariances using a Bonferroni correction ($.05/10 = .005$) we found no statistically significant differences between the pre-IB and IB groups on the factor covariances (all $ps > .02$).

In summarizing the tests of measurement invariance, we determined that at the item level the unstandardized factor loadings did not differ significantly between the pre-IB and IB groups.

We identified three items within the Motivation/Self-Regulation factor where the intercepts for the pre-IB group were significantly higher than those from the IB group. Variability in the Attitudes Toward School factor was greater for the IB group.

Cronbach's Alpha Reliability

Study 1. Table 6 contains Cronbach's alpha reliabilities for the pre-IB and IB groups for Studies 1 and 2, along with the values reported by McCoach and Siegle (2003a) for their sample of 645 high school students of varying achievement levels. Reliabilities for the scores for the five factors in Study 1 were similar for the pre-IB and IB groups and ranged from .83 (Attitudes Toward Teachers for the IB group) to .93 (Attitudes Toward School for the IB group). McCoach and Siegle's reported reliabilities ranged from .86 to .91.

Study 2. Cronbach's alpha reliabilities in Study 2 were generally higher than those in Study 1 and ranged from .87 (Academic Self-Perceptions) to .93 (Attitudes Toward School) for the pre-IB groups and from .87 (Academic Self-Perceptions) to .94 (Attitudes Toward School) for the IB group.

Relationships Between the Five Factors from the SAAS-R and Students' Grade Point Average

Study 1. Structural equation modeling (SEM) was used to examine the relationships between the five SAAS-R factors (predictor variables) and the outcome variable of students' cumulative grade point averages (GPA) obtained from students' records. These relationships, which were used to evaluate the criterion-related validity of the SAAS-R scores, were obtained by adding GPA to the five-factor modified CFA model (i.e., included three pairs of error covariances). Pre-IB students' GPAs ranged from 2.38 to 4.00 ($M = 3.57$, $SD = 0.40$). In

interpreting the results of these models it is important to keep in mind that the confirmatory factor analysis models in Study 1 had less than acceptable fit.

For the pre-IB group, Academic Self-Perceptions ($r = .33$), Goal Valuation ($r = .32$), and Motivation/Self-Regulation ($r = .41$) had statistically significant ($p < .001$) positive correlations with GPA. The five factors explained 23.2% ($p < .001$) of the variance in students' GPAs. When controlling for the other SAAS-R factors in the model, the unstandardized regression coefficients for Academic Self-Perceptions ($b = 0.13$, $p < .05$) and Motivation/Self-Regulation ($b = 0.13$, $p < .05$) were statistically significant (see Table 7).

IB students' GPAs ranged from 2.47 to 4.00 ($M = 3.53$, $SD = 0.37$). For the IB students, all five SAAS-R factors had statistically significant positive correlations with GPA. The five factors explained 23.9% ($p < .001$) of the variance in students' GPAs. After controlling for the other SAAS-R factors in the model, the unstandardized regression coefficients for Academic Self-Perceptions ($b = 0.19$, $p < .01$) and Motivation/Self-Regulation ($b = 0.10$, $p < .01$) were statistically significant (see Table 7).

As an additional analysis to evaluate the criterion-related validity of the SAAS-R scores, we followed the approach used by McCoach and Siegle (2003a), which involved comparing the observed scores on the SAAS-R factors for academically able achievers and underachievers. Able achievers in Study 1 were defined using McCoach and Siegle's criterion as those students with a GPA of 3.75 or higher, and underachievers were defined as those with a GPA of 2.99 or lower (McCoach and Siegle [2003a] used 2.5 or lower in their study for underachievers but because of the smaller sample sizes in Study 1 we used a less stringent cutoff to have a sufficient number of students in the underachiever group). Independent t -tests for the pre-IB students revealed statistically significant differences between the groups for the Academic Self-

Perceptions, Goal Valuation, and Motivation/Self-Regulation factors (see Table 8 for results of the independent *t*-tests along with the values for Cohen's *d* effect size). For the IB students, the two achievement groups differed significantly on all factors except for Attitudes Toward School with effect sizes ranging from 0.35 (Attitudes Toward School) to 1.52 (Academic Self-Perceptions). The average of the five effect sizes for the pre-IB group ($M = 0.53$, $SD = 0.40$, range = -0.02 to 0.98) was lower than McCoach and Siegle's (2003a) average effect size ($M = 0.89$, $SD = 0.36$, range = 0.41 to 1.57); the IB group's average effect size was slightly higher ($M = 0.93$, $SD = 0.47$) than the average of McCoach and Siegle's (2003a) effects.

Study 2. As in Study 1, GPA was added to the five-factor confirmatory factor analysis model (three error covariances were included) for the SAAS-R. Pre-IB students' semester GPAs ranged from 0.33 to 4.00 ($M = 3.32$, $SD = 0.65$) and IB students' semester GPAs ranged from 1.00 to 4.00 ($M = 3.28$, $SD = 0.56$). Correlations between the five SAAS-R factors and students' GPAs were statistically significant ($ps < .001$) for all factors for both pre-IB and IB students. Attitudes Toward School showed the weakest relation to students' GPAs for the pre-IB and IB students ($rs = .16$ and $.15$, respectively). Motivation/Self-Regulation had the strongest relation to students' GPAs for both groups ($rs = .42$ and $.41$, respectively). In the structural equation model, the five factors explained 19.3% and 19.4% ($p < .001$) of the variance in pre-IB and IB students' GPAs, respectively. After controlling for the other SAAS-R factors in the model, the unstandardized regression coefficients for Academic Self-Perceptions and Motivation/Self-Regulation factors were statistically significant (see Table 7).

Similar to Study 1, we used the approach by McCoach and Siegle (2003a) to compare the high achievers with the underachievers as a source of criterion-related validity evidence for the five SAAS-R scores. As in Study 1 high achievers were defined as those students with GPAs of

3.75 or higher, but underachievers were defined as those with a GPA of 2.50 or lower (McCoach and Siegle [2003a] used 2.5 or lower in their study). Independent *t*-tests for the pre-IB students and IB groups revealed statistically significant differences between the groups for all five factors (see Table 9 for results of the independent *t*-tests along with the values for Cohen's *d* effect size). The averages of the five effect sizes for the pre-IB ($M = 0.89$, $SD = 0.31$, range = 0.60 to 1.32) and IB groups ($M = 0.92$, $SD = 0.44$) were similar to McCoach and Siegle's (2003a) average effect ($M = 0.89$, $SD = 0.36$, range = 0.41 to 1.57).

Discussion

As researchers continue to explore the psychological factors associated with students' academic achievement, there is a need for measurement instruments that produce reliable and valid scores for various student populations including those in specialized programs such as the International Baccalaureate (IB) Diploma program. The SAAS-R represents an important addition to the measurement tools available to researchers and its importance is reflected in its increasing use in the field. Despite this increasing use there has not been a corresponding increase in researchers' efforts to evaluate the psychometric properties of the SAAS-R when used with their own data. Given that problems with score reliability and validity can compromise statistical analyses such as multiple regression and ANOVA (Vacha-Haase & Thompson, 2011), there is a need to better understand and evaluate the psychometric properties of measures like the SAAS-R. To address this need, we evaluated the reliability, factor structure, measurement invariance, and criterion-related validity of the scores from the SAAS-R for two samples of students in pre-IB and IB programs. In interpreting the results of the present two studies, it is important to note that the schools and students were from one southeastern state and were not randomly selected. Only students who had parent permission and who completed signed assent

forms participated. Data also were collected at one point in time and, therefore, it is not possible to evaluate the stability of the results over time.

Overall the reliabilities of the scores of the five SAAS-R factors were good, with all values from the pre-IB and IB groups across the two studies greater than .80. The importance of estimating reliability for one's own scores and reporting these values, rather than relying on the values in a test manual, is underscored by the finding of considerable variation in the reliability estimates across the various groups and conditions in the current two studies. The greatest variation in the reliability estimates across the subscales was for Attitudes Toward Teachers with Cronbach's alphas ranging from .83 (IB in Study 1) to .92 (pre-IB and IB groups in Study 2). This is the only subscale that was modified in Study 2 based on students' feedback from Study 1 that there was some confusion about the frame of reference for answering questions about their teachers. The modification involved clarifying that student respondents should consider their *IB* teachers and their *IB* classes when answering the seven questions within the Attitudes Toward Teachers subscale (e.g., I relate well to my *IB* teachers). Clarifying the frame of reference for answering the SAAS-R items (the first item on the SAAS-R was one of the seven items that was changed and may have provided the proper frame of reference for *all* of the SAAS-R items) may have resulted in the improved reliabilities and model fit exhibited in Study 2. It should be noted that in addition to changing the wording for seven items in Study 2, the sample was much larger in Study 2 ($n = 1149$ vs. $n = 312$) and came from a more heterogeneous group of schools; the diversity of schools in Study 2 resulted in slightly more variability in the students' demographic characteristics (see Table 1). Additional research is needed to determine if this improved reliability and model fit with the wording change are replicated with different samples of IB students. These results also suggest that ongoing efforts to use qualitative approaches like

cognitive interviewing (Willis, 1999) are recommended to understand how students comprehend and interpret the content of questionnaire items.

The variability in the reliability estimates for the five subscales is even greater when one considers the values from studies that have used the SAAS-R (see online supplemental materials, Appendix A; e.g., values for the Attitudes Toward Teachers ranged from .68 to .92). Reporting information about score reliability is critical for interpreting the results of statistical analyses (e.g., low reliability may attenuate statistical relations between variables). The reported reliabilities also can be used by researchers conducting reliability generalization studies that are designed to describe and explain the variability in score reliability across groups and conditions with the goal of improving the measurement process. In the present review of studies that used the SAAS-R, 63% calculated Cronbach's alpha reliabilities for the data collected in their studies, which was higher than what Warne et al. (2012) found (53.3%) in articles published recently in gifted education journals.

While information about the reliability of scores from an instrument is important, this information does not test whether the factor structure underlying the instrument is consistent with newly collected data. Factor structures may change over time and vary across different subgroups of students, and therefore it is critical to evaluate the extent to which the measurement model underlying the instrument fits the data that have been collected. To date, there have been a limited number of factor analyses of the SAAS-R. In Study 1, the fit indices of the five-factor model for the SAAS-R were generally below the acceptable values, suggesting greater misfit than McCoach and Siegle's (2003a) model. In contrast, in Study 2 the fit of the same five-factor model to the data from a slightly modified version of the SAAS-R for the pre-IB and IB students was similar to McCoach and Siegle's (2003a) model. The variation in model fit when an

instrument is used with different samples under different conditions underscores the importance of conducting and reporting the results of CFAs so that researchers have a better sense of the robustness of the factor structure and can gauge the meaningfulness of the statistical comparisons using the scores from the instrument. In addition, reporting the results of CFAs can provide the original test developers and other researchers with information about the strengths and weaknesses of the measure and potential ways of improving the measure. In the present study, there were areas of model misfit that were consistent across groups and studies. For example, there was misfit that involved covariation of the errors of similarly-worded pairs of items (e.g., 2. “I am intelligent”/5. “I am smart in school”; 30. “I spend a lot of time on my schoolwork”/33. “I put a lot of effort into my schoolwork”; 1. “My classes are interesting” /34. “I like my classes”). If these sources of misfit (covariances for these three pairs of items) are replicated in future research, one solution would be to remove one of the two similarly worded items in each pair. For exploratory purposes we removed one item from each pair of items that exhibited correlated errors to determine the effect on reliability and model fit for the Study 2 data. Based on consultations with doctoral students in measurement and psychology and discussion within the research team, we decided to remove items 1, 2, and 30. The largest drop in reliability was for Academic Self-Perceptions for the IB group, which decreased from .87 to .84. Overall model fit was acceptable for the pre-IB and IB groups, respectively: $SB\chi^2(454, N = 589) = 1115.32, p < .0001$ (CFI = .92, TLI = .92, RMSEA = .05, SRMR = .05) and $SB\chi^2(454, N = 560) = 1057.24, p < .0001$ (CFI = .93, TLI = .93, RMSEA = .05, SRMR = .05). These results suggest that it would be reasonable to drop these three items if the measure was used with IB students. If researchers decide not to drop these items it is important that they examine the SAAS-R using a latent variable framework (rather than observed variables as reflected in subscale scores) so that

measurement error and correlated errors can be taken into account when examining relationships between the SAAS-R scores and relevant outcomes.

As a relatively new instrument the equivalence (invariance) of the psychometric properties of the SAAS-R (i.e., item factor loadings, item intercepts, residuals for items) has not been examined for different groups of students. The present study's tests of measurement invariance served two purposes. First, these tests evaluated if the psychometric properties of the SAAS-R items were functioning similarly across the pre-IB and IB groups as a way of evaluating potential biases in the instrument's scores. Secondly, these invariance tests provided another approach to exploring how two groups of students (pre-IB and IB) at different developmental stages (9th – 12th grade) and with different curricular experiences perceive themselves, their teachers, and schools. Overall, the results from the measurement invariance tests showed that the structure of students' attitudes and the relationship of the items to the five factors underlying the SAAS-R were similar for pre-IB and IB students. The biggest difference between the groups showed up on the intercepts of three items in the Motivation/Self-Regulation factor where, after controlling for students' level of Motivation/Self-Regulation, pre-IB students more strongly endorsed these items (i.e., uniform DIF). These items ("I check my assignments before I turn them in"; "I complete my schoolwork regularly"; "I am organized about my schoolwork") may represent a secondary dimension or factor within the Motivation/Self-Regulation factor that reflects student consciousness. While there is little agreement on what steps should be taken if non-invariance is detected, one option for handling DIF is to delete the non-invariant items from the analyses (Cheung & Rensvold, 1999; Millsap, 2011; Sass, 2011). When these three items were deleted, the pre-IB group was 0.15 standard deviations higher on the latent mean for Motivation/Self-Regulation compared to the IB group. A second option for

handling DIF is to retain these items and use caution in comparing the latent means. In the present study, the differences in the item intercepts were relatively small so the intercepts were retained and allowed to be freely estimated in each group. Retaining these three items and comparing the latent means for Motivation/Self-Regulation produced nearly identical results. Additional measurement invariance studies are needed to determine if these results will be replicated.

Although the value of measurement invariance studies is widely recognized, there are challenges associated with implementing these studies. With multiple groups, larger sample sizes are needed and with instruments like the SAAS-R that contain many items measuring several factors, the number and complexity of statistical comparisons increase substantially. The fact that there is not one universally accepted approach to conducting invariance tests or for evaluating the practical significance of differences on the various item characteristics (e.g., intercepts) adds to the complexity.

Finally, support for the criterion-related validity of the scores from the SAAS-R subscales was provided by the significant, positive relationships between these factors and students' GPAs. Consistent with McCoach and Siegle's (2003a) finding of a large effect for the Motivation/Self-Regulation subscale between gifted high achievers and underachievers, the present two studies found that one of the strongest relations was between this subscale and students' GPAs. Academic Self-Perceptions and Goal Valuation, two additional self-directed beliefs reflecting students' levels of cognitive engagement with school, also showed moderate to strong relations with students' GPAs. The SAAS-R factors assessing more other-directed beliefs reflecting affective engagement, Attitudes Toward School and Attitudes Toward Teachers, tended to yield smaller associations.

Overall the results from the current two studies highlight the importance of evaluating the psychometric properties of a measure each time it is administered. Results from Study 2 focusing on the reliability and validity of scores from a modified SAAS-R provide strong support for the use of the SAAS-R as a research tool to examine the psychological factors associated with pre-IB and IB students' academic achievement. Results from Study 1, which were based on a smaller and more homogeneous sample of pre-IB and IB students, were not as strong in terms of reliability and model fit. Additional research, using national samples with larger and more diverse groups of pre-IB and IB students, is needed to determine if these results will be replicated. Ongoing validation work also is needed to evaluate how the five factors underlying the SAAS-R relate to other theoretically relevant variables such as students' academic engagement and self-efficacy. For researchers using the SAAS-R with other populations and under different conditions it is critical that reliability and validity of the scores are evaluated for the researchers' own data and reported in the literature so that variation in the psychometric properties of the SAAS-R can be described, and factors associated with this variation (e.g., sampling error) can be explored.

References

- American Educational Research Association, American Psychological Association, and the National Council on Measurement in Education (2014). *Standards for educational and psychological testing*. Washington, DC: AERA.
- Bentler, P. M. (1992). On the fit of models to covariances and methodology to the *Bulletin of Psychological Bulletin*, *112*, 400–404. doi: 10.1037/0033-2909.112.3.400
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*, *14*, 464-504. doi: 10.1080/10705510701301834
- Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., & Paxton, P. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociological Methods & Research*, *36*, 462-494. doi: 10.1177/0049124108314720
- Cheung, G. W., & Rensvold, R. B. (1999). Testing factorial invariance across groups: A reconceptualization and proposed new method. *Journal of Management*, *25*(1), 1-27. doi: 10.1177/014920639902500101
- Colangelo, N., Assouline, S. A., & Gross, M. U. M. (Eds.) (2004). *A nation deceived: How schools hold back America's brightest students*. The Templeton National Report on Acceleration. Iowa City, IA.
- Davie, J. L. (2012). *Giftedness and underachievement: A comparison of student groups* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No.3513325)
- Dimitrov, D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development*, *43*, 121-149. doi:

0.1177/0748175610373459

DiStefano, C., & Hess, B. (2005). Using confirmatory factor analysis for construct validation:

An empirical review. *Journal of Psychoeducational Assessment*, 23, 225-241. doi:

10.1177/073428290502300303

Duran, R. P., Eisenhart, M. A., Erickson, F. D., Grant, C. A., Green, J. L., Hedges, L. V., . . .

Schneider, B. L. (2006). Standards for reporting on empirical social science research in

AERA publications: American Educational Research Association. *Educational*

Researcher, 35(6), 33–40. doi: 10.3102/0013189X035006033

Hertberg-Davis, H., & Callahan, C. M. (2008). A narrow escape: Gifted students perceptions of

Advanced Placement and International Baccalaureate Programs. *Gifted Child Quarterly*,

52, 199-216. doi: 10.1177/0016986208319705

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indices in covariance structure analysis:

Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6, 1–55.

doi: 10.1080/10705519909540118

International Baccalaureate Organization (2013a). *History of International Baccalaureate*.

Retrieved from <http://www.ibo.org/history/>

International Baccalaureate Organization (2013b). *IB world school statistics*. Retrieved from

<http://www.ibo.org/facts/schoolstats/progsbycountry.cfm>

International Baccalaureate Organization (2013c). *Information for United States*. Retrieved from

<http://www.ibo.org/country/US/>

International Baccalaureate Organization (2013d). *Programme growth statistics*. Retrieved from

<http://www.ibo.org/facts/schoolstats/growth.cfm>

International Baccalaureate Organization (2013e). *The IB diploma programme*. Retrieved from

<http://www.ibo.org/diploma/>

International Baccalaureate Organization (2013f). *The Middle Years Programme*. Retrieved from

<http://www.ibo.org/myp/>

Kieffer, K. M., Reese, R. J., & Vacha-Haase, T. (2010). Reliability generalization methods in the context of giftedness research. In B. Thompson & R. Subotnik & (Eds.), *Methodologies for conducting research on giftedness* (pp. 89-111). Washington, DC: American Psychological Association. doi: 10.1037/12079-005

Psychological Association. doi: 10.1037/12079-005

Matthews, M. S., & McBee, M. T. (2007). School factors and the underachievement of gifted students in a talent search summer program. *Gifted Child Quarterly, 51*, 167-181. doi:

10.1177/0016986207299473

McCoach, D. B. (2002). A validation study of the School Attitude Assessment Survey.

Measurement and Evaluation in Counseling and Development, 35, 66-77.

McCoach, D. B., & Siegle, D. (2003a). The School Attitude Assessment Survey-Revised: A new instrument to identify academically able students who underachieve. *Educational and Psychological Measurement, 63*, 414-429. doi: 10.1177/0013164403063003005

Psychological Measurement, 63, 414-429. doi: 10.1177/0013164403063003005

McCoach, D. B., & Siegle, D. (2003b). The structure and function of academic self-concept in gifted and general education students. *Roepers Review, 25*(2), 61-65. doi:

10.1080/02783190309554200

Millsap, R. E. (2011). *Statistical approaches to measurement invariance*. New York, NY:

Routledge.

Muthén, L. K., & Muthén, B. O. (1998-2012). *Mplus user's guide*. Seventh Edition.

Los Angeles, CA: Muthén & Muthén.

- Perna, L. W., May, H., Yee, A., Ransom, T., Rodriguez, A., & Fester, R. (2011). The potential role of the International Baccalaureate (IB) Diploma Programme in improving academic preparation for college for all students. The Consortium for Policy Research in Education. Retrieved from <http://www.cpre.org/ib>
- Peters, S. J., & Gentry, M. (2010). Multigroup construct validity evidence of the HOPE Scale: Instrumentation to identify low-income elementary students for gifted programs. *Gifted Child Quarterly, 54*, 298-313. doi: 10.1177/0016986210378332
- Sass, D. (2011). Testing measurement invariance and comparing latent factor means within a confirmatory factor analysis framework. *Journal of Psychoeducational Assessment, 29*, 347-363. doi: 10.1177/0734282911406661
- Suldo, S. M., Shaffer, E. J., & Shaunessy, E. (2008). An independent investigation of the validity of the School Attitude Assessment Survey-Revised. *Journal of Psychoeducational Assessment, 26*, 69-82. doi:10.1177/0734282907303089
- Sun, J. (2005). Assessing goodness of fit in confirmatory factor analysis. *Measurement and Evaluation in Counseling and Development, 37*, 240-256.
- Thompson, B., & Vacha-Haase, T. (2000). Psychometrics is datametrics: The test is not reliable. *Educational and Psychological Measurement, 60*, 174-195. doi: 10.1177/0013164400602002
- Vacha-Haase, T., & Thompson, B. (2011). Score reliability: A retrospective look back at 12 years of reliability generalization studies. *Measurement and Evaluation in Counseling and Development, 44*, 159-168. doi: 10.1177/0748175611409845

Warne, R. T., Lazo, M., Ramos, T., & Ritter, N. (2012). Statistical methods used in gifted education journals, 2006-2010. *Gifted Child Quarterly*, 56. 134-149.

doi:10.1177/0016986212444122

Willis, G. B. (1999). *Cognitive interviewing: A "how-to" guide*. Retrieved from

<http://fog.its.uiowa.edu/~c07b209/interview.pdf>

Table 1

Demographic Characteristics of the High School Students in Pre-IB and IB (International Baccalaureate) Programs for Studies 1 and 2

Characteristic	Study 1				Study 2			
	Pre-IB (n = 161)		IB (n = 151)		Pre-IB (n = 589)		IB (n = 560)	
	n	%	n	%	n	%	n	%
Gender								
Male	61	37.9	60	39.7	244	41.4	222	39.6
Female	100	62.1	91	60.3	345	58.6	338	60.4
Free or Reduced Lunch								
No	138	85.7	135	89.4	447	75.9	423	75.5
Yes	23	14.3	15	9.9	142	24.1	137	24.5
Missing	0	0.0	1	0.7	0	0.0	0	0.0
Race								
White	79	49.1	92	60.9	298	50.6	267	47.7
African American	6	3.7	9	6.0	71	12.1	91	16.3
American Indian	3	1.9	0	0.0	1	0.2	1	0.2
Asian	43	26.7	33	21.9	111	18.8	113	20.2
Native Hawaiian, Other Pacific Islander	0	0.0	0	0.0	1	0.2	1	0.2
Other	14	8.7	5	3.3	42	7.1	37	6.6
Multiple Categories	13	8.1	12	7.9	65	11.0	50	8.9
Missing	3	1.9	0	0.0	0	0.0	0	0.0
Hispanic, Latino, or Spanish Origin								
No	137	85.1	139	92.1	501	85.1	472	84.3
Mexican	2	1.2	1	0.7	7	1.2	7	1.3
Puerto Rican	7	4.3	3	2.0	11	1.9	11	2.0
Cuban	2	1.2	2	1.3	10	1.7	4	0.7
Other	13	8.1	6	4.0	60	10.2	66	11.8
Father's Education								
8th Grade or Less	3	1.9	0	0.0	1	0.2	10	1.8
Some H.S.	6	3.7	4	2.6	25	4.2	27	4.8
H.S. Diploma/ GED	14	8.7	19	12.6	101	17.1	87	15.5
Some College	11	6.8	20	13.2	68	11.5	76	13.6

Characteristic	Study 1				Study 2			
	Pre-IB (<i>n</i> = 161)		IB (<i>n</i> = 151)		Pre-IB (<i>n</i> = 589)		IB (<i>n</i> = 560)	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
College/University	61	37.9	60	39.7	197	33.4	188	33.6
Master's	34	21.1	28	18.5	119	20.2	107	19.1
Beyond Master's	30	18.6	20	13.2	60	10.2	60	10.7
Missing	2	1.2	0	0.0	18	3.1	5	0.8
Mother's Education								
8th Grade or Less	0	0.0	4	2.6	4	0.7	2	0.4
Some H.S.	6	3.7	4	2.6	14	2.4	14	2.5
H.S. Diploma/ GED	17	10.6	17	11.3	69	11.7	74	13.2
Some College	14	8.7	21	13.9	72	12.2	82	14.6
College/University	75	46.6	65	43.0	246	41.8	260	46.4
Master's	32	19.9	32	21.2	138	23.4	93	16.6
Beyond Master's	16	9.9	8	5.3	41	7.0	34	6.1
Missing	1	0.6	0	0.0	5	0.8	1	0.2
Parent Status								
Married	121	75.2	114	75.5	432	73.3	381	68.0
Divorced	25	15.5	29	19.2	95	16.1	111	19.8
Separated	4	2.5	0	0.0	12	2.0	13	2.3
Never Married	5	3.1	4	2.6	35	5.9	42	7.5
Never Married But Living Together	0	0.0	0	0.0	5	0.8	6	1.1
Widowed	5	3.1	4	2.6	10	1.7	7	1.2
Missing	1	0.6	0	0.0	0	0.0	0	0.0

Note. Percentages may not sum to 100 due to rounding. For Studies 1 and 2, pre-IB and IB students did not differ significantly on the demographic characteristics included in the table based on the results of chi-square analyses ($p > .01$).

Table 2

Fit Indices for the School Attitude Assessment Survey-Revised Using Robust Maximum Likelihood Estimation by Study and International Baccalaureate (IB) Group Along with Values Reported by McCoach and Siegle (2003a) and Davie (2012)

	Study 1		Study 2		McCoach & Siegle (2003a)	Davie (2012)
	Pre-IB	IB	Pre-IB	IB		
<i>n</i>	161	151	589	560	537	847
χ^2	984.42	1043.84	1451.15	1426.05	1581.70	1634.17
<i>df</i>	550	550	550	550	550	550
CFI	.86	.83	.91	.92	.91	.94
TLI	.84	.82	.90	.91	.92	.94
RMSEA	.07 (.06, .08) ^a	.08 (.07, .08) ^a	.05 (.05, .06) ^a	.05 (.05, .06) ^a	.06 ^b	.05 ^b
SRMR	.08	.09	.05	.06	.06	.04

Note. CFI = Comparative Fit Index; TLI= Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation; SRMR = Standardized Root Mean Square Residual.

^a Numbers in parentheses represent the 90% confidence interval for the RMSEA.

^b Confidence interval not reported for the RMSEA.

Table 3

Standardized Factor Pattern Coefficients (Loadings) for the Five-Factor School Attitude Assessment Survey-Revised (SAAS-R) Using Robust Maximum Likelihood Estimation by International Baccalaureate (IB) Group and Study Along with Values Reported by McCoach and Siegle (2003a)

Factor	No. of Items	Study 1 Standardized Factor Loadings				Study 2 Standardized Factor Loadings				McCoach & Siegle (2003a; $n = 537$)	
		Pre-IB ($n = 161$)		IB ($n = 151$)		Pre-IB ($n = 589$)		IB ($n = 560$)			
		$M (SD)$	Range	$M (SD)$	Range	$M (SD)$	Range	$M (SD)$	Range	$M (SD)$	Range
ASP	7	.74 (.11)	.55 to .85	.73 (.11)	.56 to .85	.73 (.11)	.53 to .84	.73 (.13)	.52 to .85	.69 (.09)	.58 to .80
ATT	7	.73 (.07)	.62 to .82	.65 (.12)	.43 to .81	.79 (.05)	.74 to .86	.79 (.04)	.74 to .85	.74 (.06)	.69 to .84
ATS	5	.81 (.05)	.77 to .91	.85 (.06)	.79 to .95	.86 (.04)	.82 to .92	.88 (.06)	.78 to .95	.79 (.12)	.58 to .88
Goals	6	.77 (.08)	.68 to .88	.73 (.10)	.56 to .84	.76 (.08)	.67 to .85	.79 (.08)	.69 to .88	.79 (.12)	.56 to .91
MOT/S -R	10	.59 (.17)	.43 to .86	.70 (.12)	.47 to .82	.69 (.13)	.50 to .87	.73 (.12)	.50 to .87	.75 (.15)	.41 to .89

Note. ASP = Academic Self-Perceptions; ATT = Attitudes Toward Teachers; ATS = Attitudes Toward School; Goals = Goal Valuation; MOT/S-R = Motivation/Self-Regulation.

Table 4

Interfactor Correlations for School Attitude Assessment Survey-Revised by Study and International Baccalaureate (IB) Group Along with Values Reported by McCoach and Siegle (2003a)

Factor	1	2	3	4
Study 1				
1. ASP				
2. ATT	.36 ^a /.41 ^b /.36 ^c			
3. ATS	.32 ^a /.25 ^b /.27 ^c	.61 ^a /.34 ^b /.60 ^c		
4. Goals	.26 ^a /.25 ^b /.37 ^c	.26 ^a /.31 ^b /.51 ^c	.31 ^a /.33 ^b /.34 ^c	
5. MOT/S-R	.35 ^a /.43 ^b /.34 ^c	.31 ^a /.36 ^b /.65 ^c	.28 ^a /.28 ^b /.45 ^c	.74 ^a /.66 ^b /.74 ^c
Study 2				
1. ASP				
2. ATT	.51 ^a /.51 ^b /.36 ^c			
3. ATS	.40 ^a /.36 ^b /.27 ^c	.60 ^a /.60 ^b /.60 ^c		
4. Goals	.34 ^a /.27 ^b /.37 ^c	.40 ^a /.33 ^b /.51 ^c	.32 ^a /.21 ^b /.34 ^c	
5. MOT/S-R	.54 ^a /.33 ^b /.34 ^c	.58 ^a /.45 ^b /.65 ^c	.42 ^a /.31 ^b /.45 ^c	.63 ^a /.68 ^b /.74 ^c

Note. ASP = Academic Self-Perceptions; ATT = Attitudes Toward Teachers; ATS = Attitudes Toward School; Goals = Goal Valuation; MOT/S-R = Motivation/Self-Regulation.

^a Pre-IB ($n = 161$ and 589 for Studies 1 and 2, respectively).

^b IB ($n = 151$ and 560 for Studies 1 and 2, respectively).

^c McCoach and Siegle (2003a); $n = 537$.

Table 5

Invariance Tests for the School Attitude Assessment Survey-Revised by IB Group (Study 2)

Model	SB χ^2	<i>df</i>	Comparison	Δ SB χ^2	Δ <i>df</i>	CFI	Δ CFI	TLI	Δ TLI	RMSEA	Δ RMSEA	SRMR	Δ SRMR
1A. Configural (same form) ^a	2877.58	1100				.91		.90		.05 (.05, .06)		.06	
1B. Configural (same form) ^b	2554.54	1094	1A vs. 1B	182.23 ^c	6	.93	.02 ^c	.92	.02 ^c	.05 (.05, .05)	.00 ^c	.05	.00 ^c
Pre-IB	1289.17	547				.92		.92		.05 (.05, .05)		.05	
IB	1264.94	547				.93		.93		.05 (.05, .05)		.06	
2. Metric (equal loadings)	2613.13	1124	2 vs. 1B	60.72 <i>p</i> < .001	30	.93	.00	.92	.00	.05 (.05, .05)	.00	.06	.01
2A. Equal Loadings: ASP	2561.80	1100	2A vs. 1B	7.86 <i>p</i> > .05	6	.93	.00	.92	.00	.05 (.05, .05)	.00	.05	.00
2B. Equal Loadings: ATT	2559.79	1100	2B vs. 1B	5.13 <i>p</i> > .05	6	.93	.00	.92	.00	.05 (.05, .050)	.00	.05	.00
2C. Equal Loadings: ATS	2573.46	1098	2C vs. 1B	20.30 <i>p</i> < .001	4	.93	.00	.92	.00	.05 (.05, .05)	.00	.05	.00
2D. Equal Loadings: Goals	2554.29	1099	2D vs. 1B	6.75 <i>p</i> > .05	5	.93	.00	.92	.00	.05 (.05, .05)	.00	.06	.01
2E. Equal Loadings: MOTS/S-R	2581.94	1103	2E vs. 1B	27.52 <i>p</i> < .01	9	.93	.00	.92	.00	.05 (.05, .05)	.00	.06	.01

Model	SB χ^2	<i>df</i>	Comparison	Δ SB χ^2	Δ <i>df</i>	CFI	Δ CFI	TLI	Δ TLI	RMSEA	Δ RMSEA	SRMR	Δ SRMR
3. Scalar (equal intercepts)	2687.93	1154	3 vs. 2	76.39 <i>p</i> < .00001	30	.92	.01	.92	.00	.05 (.05, .05)	.00	.06	.00
3A. Equal Intercepts: ASP	2626.55	1130	3A vs. 2	13.26 <i>p</i> < .05	6	.93	.00	.92	.00	.05 (.05, .05)	.00	.06	.00
3B. Equal Intercepts: ATT	2625.41	1130	2B vs. 2	11.79 <i>p</i> > .05	6	.93	.00	.92	.00	.05 (.05, .050)	.00	.06	.00
3C. Equal Intercepts: ATS	2620.65	1128	2C vs. 2	6.98 <i>p</i> > .05	4	.93	.00	.92	.00	.05 (.05, .05)	.00	.06	.00
3D. Equal Intercepts: Goals	2623.40	1129	2D vs. 2	9.75 <i>p</i> > .05	5	.93	.00	.92	.00	.05 (.05, .05)	.00	.06	.00
3E. Equal Intercepts: MOTS/S-R	2644.41	1133	2E vs. 2	34.71 <i>p</i> < .001	9	.93	.00	.92	.00	.05 (.05, .05)	.00	.06	.00
3.1 Scalar (equal intercepts except for items 4, 24, 26)	2673.45	1151	3.1 vs. 2	59.54 <i>p</i> < .001	27	.92	.01	.92	.00	.05 (.05, .050)	.00	.06	.00
4. Equal Item Residual (Error) Variances	2697.33	1186	4 vs. 3.1	49.80 <i>p</i> > .05	35	.93	.01 _c	.92	.00	.05 (.05, .05)	.00	.07	.01
4A. Equal Error Covariances _b	2692.69	1189	4A vs. 4	1.64 <i>p</i> > .05	3	.93	.00	.93	.00 _c	.05 (.05, .05)	.00	.07	.00

Model	SB χ^2	<i>df</i>	Comparison	Δ SB χ^2	Δ <i>df</i>	CFI	Δ CFI	TLI	Δ TLI	RMSEA	Δ RMSEA	SRMR	Δ SRMR
5. Equal factor variances	2709.22	1194	5 vs. 4A	15.56 (<i>p</i> < .01)	5	.92	.01	.93	.00	.05 (.05, .05)	.00	.08	.01
6. Equal covariances between factors	2719.41	1199	6 vs. 4A	26.39 (<i>p</i> < .01)	10	.92	.01	.93	.00	.05 (.05, .05)	.00	.08	.01

Note. ASP = Academic Self-Perceptions; ATT = Attitudes Toward Teachers; ATS = Attitudes Toward School; Goals = Goal Valuation; MOT/S-R = Motivation/Self-Regulation. SB χ^2 = Satorra-Bentler scaled chi-square.

*** *p* < .001.

^a See Table 2 for fit statistics for the pre-IB and IB groups.

^b 3 pairs of error covariances (2 and 5; 1 and 34; 30 and 33) were added to the model within each group.

^c Change in fit represents improved fit.

Table 6

Cronbach's Alpha Reliabilities for the Five Factors of the School Attitude Assessment Survey-Revised (SAAS-R) by Study and International Baccalaureate (IB) Group Along with Values Reported by McCoach and Siegle (2003a)

Factor	No. of Items	Study 1				Study 2				McCoach & Siegle (2003a)
		Pre-IB (n = 161)		IB (n = 151)		Pre-IB (n = 589)		IB (n = 560)		
		Cronbach's Alpha	Item-to-Total Correlation Range	Cronbach's Alpha	Item-to-Total Correlation Range	Cronbach's Alpha	Item-to-Total Correlation Range	Cronbach's Alpha	Item-to-Total Correlation Range	
ASP	7	.88	.53-.77	.87	.58-.76	.87	.51-.76	.87	.52-.74	.86
ATT	7	.89	.58-.76	.83	.38-.74	.92	.70-.83	.92	.70-.81	.89
ATS	5	.90	.71-.85	.93	.76-.90	.93	.79-.88	.94	.77-.91	.87
Goals	6	.90	.65-.82	.87	.54-.75	.89	.63-.79	.91	.66-.83	.89
MOT/S-R	10	.89	.48-.79	.90	.47-.76	.90	.51-.80	.92	.51-.82	.91

Note. ASP = Academic Self-Perceptions; ATT = Attitudes Toward Teachers; ATS = Attitudes Toward School; Goals = Goal Valuation; MOT/S-R = Motivation/Self-Regulation.

Table 7

Structural Equation Modeling Results of the Five Factors of the School Attitude Assessment Survey-Revised (SAAS-R) Predicting Students' GPAs by Study and International Baccalaureate Group

Study 1						
Predictor	Pre-IB (<i>n</i> = 161)			IB (<i>n</i> = 151)		
	<i>r</i>	<i>b</i>	<i>SE</i>	<i>r</i>	<i>b</i>	<i>SE</i>
ASP	.33***	0.13*	0.05	.43***	0.19**	0.07
ATT	.05	-0.09	0.05	.21**	-0.01	0.05
ATS	.10	0.01	0.04	.17*	0.01	0.02
Goals	.32***	0.04	0.10	.26***	0.02	0.08
MOT/S-R	.41***	0.13*	0.06	.38***	0.10*	0.05
<i>R</i> ₂	.23** (<i>SE</i> = .08)			.24*** (<i>SE</i> = .07)		
Study 2						
Predictor	Pre-IB (<i>n</i> = 589)			IB (<i>n</i> = 560)		
	<i>r</i>	<i>b</i>	<i>SE</i>	<i>r</i>	<i>b</i>	<i>SE</i>
ASP	.32***	0.14**	0.05	.27***	0.14***	0.04
ATT	.22***	-0.03	0.04	.19***	-0.05	0.03
ATS	.16***	-0.02	0.03	.15***	0.01	0.02
Goals	.22***	-0.07	0.08	.29***	0.02	0.07
MOT/S-R	.42***	0.31***	0.06	.41***	0.23***	0.04
<i>R</i> ₂	.19*** (<i>SE</i> = 0.03)			.19*** (<i>SE</i> = 0.04)		

Note. ASP = Academic Self-Perceptions; ATT = Attitudes Toward Teachers; ATS = Attitudes Toward School; Goals = Goal Valuation; MOT/S-R = Motivation/Self-Regulation. *b* = unstandardized regression coefficient. *SE* = standard error for the unstandardized regression coefficient. All models included three pairs of error covariances (2 and 5; 1 and 34; 30 and 33). * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 8

Effect Sizes and Independent Samples t-tests for the Five Factors from the School Attitude Assessment Survey-Revised (SAAS-R) Between High Achievers and Underachievers for Pre-IB and IB Students (Study 1)

Subscale	Pre-IB				<i>p</i>	<i>d</i>
	High Achievers (<i>n</i> = 69)		Underachievers (<i>n</i> = 16)			
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
1. ASP	5.69	0.81	4.80	1.25	.014	0.98
2. ATT	4.77	1.07	4.80	0.84	.931	-0.02
3. ATS	5.55	1.24	5.22	1.20	.334	0.27
4. Goals	6.71	0.54	6.33	0.70	.020	0.66
5. MOT/S-R	5.66	0.88	4.95	1.04	.006	0.78

Subscale	IB				<i>p</i>	<i>d</i>
	High Achievers (<i>n</i> = 50)		Underachievers (<i>n</i> = 14)			
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
1. ASP	5.95	0.67	4.73	1.15	.002	1.52
2. ATT	5.26	0.76	4.80	0.56	.038	0.64
3. ATS	5.16	1.23	4.71	1.43	.247	0.35
4. Goals	6.74	0.42	6.36	0.52	.006	0.87
5. MOT/S-R	5.68	0.78	4.62	1.05	< .001	1.26

Note. ASP = Academic Self-Perceptions; ATT = Attitudes Toward Teachers; ATS = Attitudes Toward School; Goals = Goal Valuation; MOT/S-R = Motivation/Self-Regulation. Response scale ranged from 1 (*strongly disagree*) to 7 (*strongly agree*). High achievement in Study 1 was defined as a GPA of 3.75 or higher; underachievement was defined as a GPA of 2.99 or lower. Effect sizes reported by McCoach and Siegle (2003a) were 0.46 (ASP), 0.78 (ATT), 0.67 (ATS), 1.23 (Goals), and 1.29 (MOT/S-R). McCoach and Siegle defined the achievers as gifted students “in the top 10% of their class or had at least a 3.75 GPA. Gifted underachievers were in the bottom half of their high school class or had a GPA at or below 2.5” (p. 425).

Table 9

Effect Sizes and Independent Samples t-tests for the Five Factors from the School Attitude Assessment Survey-Revised (SAAS-R) Between High Achievers and Underachievers for Pre-IB and IB Students (Study 2)

Subscale	Pre-IB				<i>p</i>	<i>d</i>
	High Achievers (<i>n</i> = 181)		Underachievers (<i>n</i> = 72)			
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
1. ASP	5.81	0.80	4.84	1.01	< .001	1.11
2. ATT	5.40	0.97	4.57	1.39	< .001	0.76
3. ATS	5.63	1.24	4.85	1.40	< .001	0.60
4. Goals	6.68	0.48	6.24	0.94	< .001	0.68
5. MOT/S-R	5.87	0.73	4.77	1.04	< .001	1.32

Subscale	IB				<i>p</i>	<i>d</i>
	High Achievers (<i>n</i> = 129)		Underachievers (<i>n</i> = 56)			
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
1. ASP	5.83	0.81	4.97	1.05	< .001	0.98
2. ATT	5.40	1.11	4.69	1.15	< .001	0.63
3. ATS	5.21	1.53	4.59	1.48	.011	0.41
4. Goals	6.69	0.48	5.91	1.16	< .001	1.03
5. MOT/S-R	5.81	0.77	4.35	1.22	< .001	1.57

Note. ASP = Academic Self-Perceptions; ATT = Attitudes Toward Teachers; ATS = Attitudes Toward School; Goals = Goal Valuation; MOT/S-R = Motivation/Self-Regulation. Response scale ranged from 1 (*strongly disagree*) to 7 (*strongly agree*). High achievement in Study 2 was defined as a GPA of 3.75 or higher; underachievement was defined as a GPA of 2.5 or lower. Effect sizes reported by McCoach and Siegle (2003a) were 0.46 (ASP), 0.78 (ATT), 0.67 (ATS), 1.23 (Goals), and 1.29 (MOT/S-R). McCoach and Siegle defined the achievers as gifted students “in the top 10% of their class or had at least a 3.75 GPA. Gifted underachievers were in the bottom half of their high school class or had a GPA at or below 2.5” (p. 425).

Appendix A

Reliability Estimates in Studies Using the School Attitude Assessment Survey-Revised Ordered by Year of Publication

Study (Year)	Sample Description	Cronbach's Alphas				
		ASP	ATT	ATS	Goals	MSR
Brandon & Lawton (2013) ^a	346 grade 3 students in the western United States.	-	-	-	-	-
Suldo, McMahan, Chappel, & Bateman (2013)	500 students in grades 9 through 12 in an urban school district in the southeastern United States.	.90	b	b	.91	b
Winton (2013)	51 intellectually gifted students in grade 10 within one midwestern school district in the United States.	-	-	-	-	-
Cruse (2012)	151 intellectually gifted girls in grades 6 through 8 from one midwestern school district in the United States.	-	-	-	-	-
Davie (2012)	847 gifted and general education students in grades 9 through 12 in a suburban Midwestern school in the United States.	.89	.87	.91	.92	.92
Figg, Rogers, McCormick, & Low (2012)	112 8 th , 9 th , and 10 th grade intellectually gifted boys attending a school for boys in Australia	-	-	-	-	-
Imasa (2012)	297 7 th and 8 th grade students attending a Midwestern school in the United States.	.93	c	c	c	c
Kirk et al. (2012)	207 8 th grade students from two urban middle schools in the United States.	d	d	d	d	d

Study (Year)	Sample Description	Cronbach's Alphas				
Rubenstein, Siegle, Reis, McCoach, & Burton (2012)	46 U. S. middle-school gifted and talented students from 19 states.	e	e	e	e	e
Bacon (2011)	101 African American students in grades 6 through 8 in multiple schools and districts within Iowa. ^f	-	-	-	-	-
Berger (2011)	19 high school students in grades 9 through 12.	-	-	-	-	-
King (2011)	127 10 th and 11 th grade students attending one of three high schools in the same district in the Southeastern United States.	.83	.89	.91	.89	.91
Robertson (2011)	280 gifted and general education students in grades 10 through 12 enrolled in advanced courses or a Governor School in the Southern United States.	-	-	-	-	-
Rubenstein (2011)	3 academically gifted 8 th grade students in a rural Northeastern school in the United States.	-	-	-	-	-
Simmons (2011)	270 African American students in grades 9 through 12 attending one school located in the Southeastern United States.	.79	.85	.91	.80	.88
Taheri (2011)	518 high school students (defined as first, second, and third grade of high school) in Iran.	.72	.87	.90	.77	.86
Tenney (2011)	182 U. S. students in grades 3, 4, and 5 from three urban Title I schools in Connecticut.	.84	.92			

Study (Year)	Sample Description	Cronbach's Alphas				
Erkman, Caner, Sart, Börkan, & Şahan, (2010)	223 5 th grade students in a private school in Turkey.	.90	.89	.94	.89	.93
Gelman (2010)	80 7 th grade students from three classes in an urban Colorado school.	-	-	-	-	-
Englund (2009)	8 White middle- and high-socioeconomic status gifted students in grades 6 through 12 in a Northern Virginia school district.	.64	.92	.91	.87	.89
Kusterer (2009)	136 16-19 year old undergraduate ($n=72$) and high school students ($n=64$) attending one university or one high school in New York.	-	-	-	-	-
Lang, Waterman, & Baker (2009)	55 adolescents (12 to 16 years old) from four affordable housing communities in Los Angeles, California.	b	b	b	.82	.92
Niederdeppe (2009)	34 gifted Latino students in grades 6 through 8 in an urban middle school in the California.	-	-	-	-	-
Rieger (2009)	1,707 11 th grade students in two districts in Western Washington state.	g	-	-	-	g
Van Eman (2009)	370 gifted students in the United States in grades 5 through 7.	.81	.92	.92	.88	.91

Study (Year)	Sample Description	Cronbach's Alphas										
Black, Little, McCoach, Purcell, & Siegle (2008)	146 6 th and 7 th grade students in the United States.	Grade 6 Time 1 = .91 Time 2 = .89	Grade 6 Time 1 = .92 Time 2 = .89	Grade 6 Time 1 = .86 Time 2 = .84	Grade 6 Time 1 = .93 Time 2 = .95							
						Grade 7 Time 1 =.91 Time 2 =.83 Time 3 =.83	Grade 7 Time 1 =.94 Time 2 =.92 Time 3 =.85	Grade 7 Time 1 =.48 Time 2 =.73 Time 3 =.80	Grade 7 Time 1 =.90 Time 2 =.85 Time 3 =.87			
						Castro (2008)	23 Latino 9 th - 11 th grade gifted students.	-	-	-	-	-
						Henry (2008)	169 high school students in the United States in grades 11 to 12 enrolled in AP or honors courses.	-	-	-	-	-
		Mason (2008)	211 10 th - 12 th grade students in the United States.	h	.68	.74	.85	.82				
		Ragland (2008)	151 high school students in grades 10 to 11.	-	-	-	-	-				
		Suldo & Shaffer (2008)	349 students in grades 6-8 in a middle school.	.90	.95	i	i	i				

Study (Year)	Sample Description	Cronbach's Alphas				
Suldo, Shaffer, & Riley (2008)	321 high school students, grades 9-12	j	.93	j	j	
Suldo, Shaffer, & Shaunessy (2008) ^k	321 high school general education and IB students in grades 9-12.	.88	-	.93	-	-
Grahek (2007)	266 undergraduate students.		.90			
Matthews & McBee (2007)	440 gifted students in the United States in grades 8 through 10.	1	1	1	1	1
Mombourquette (2007)	424 students from Alberta, Canada, grades 10-12.			c		
Baslanti & McCoach (2006)	163 undergraduate students in Turkey.	.81	.87	.90	.91	.95
Olivarez (2004)	44 Hispanic gifted students in the United States in grade 8.	.71	.92	.93	.86	.92

Note. ASP = Academic Self-Perceptions; ATT = Attitudes Toward Teachers; ATS = Attitudes Toward School; Goals = Goal Valuation; MOT/S-R = Motivation/Self-Regulation.

^aAuthors reported Cronbach's alpha reliability of .91 for a total score for a revised 26-item version of the SAAS-R.

^b Study used only two of the five subscales.

^c Study used only one of the five subscales.

^d Cronbach's alpha reliabilities reported to be > .70.

^e Cronbach's alpha reliabilities reported to be > .85.

- f Author combined the Peer Perceptions Scale from the original SAAS with the SAAS-R to create a 40 item modified SAAS/SAAS-R. Although descriptive statistics for the six subscales were presented only one Cronbach's alpha was presented (.93) for the SAAS/SAAS-R.
- g Study used two of the five subscales but did not report reliability.
- h Study used only four of the five subscales.
- i Cronbach's alpha reliabilities reported to range from .90 (Academic Self-Perceptions) to .95 (Attitudes Toward School).
- j Authors reported a Cronbach's alpha reliability of .94 for a combined score of the ASP, MSR, and Goal items.
- k Authors reported that the Cronbach's alphas ranged from .88 (Academic Self-Perceptions) to .93 (Attitudes Toward School).
- l Authors indicate that reliability estimates on all 5 of the SAAS-R subscales are comparable to McCoach and Siegle (2003a).

References for Appendix A

(Online Supplemental Material)

- Bacon, L. C. (2011). *Academic self-concept and academic achievement of African American students transitioning from urban to rural schools* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 894472427)
- Baslanti, U., & McCoach, D. B. (2006). Factors related to the underachievement of university students in Turkey. *Roeper Review*, 28, 210-215. doi: 10.1080/02783190609554366
- Berger, K. C. (2011). *A research utilization project: Implementation of an evidence-based behavioral treatment for students at-risk of dropout at Richland Northeast High School* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3454672)
- Black, A. C., Little, C. A., McCoach, D. B., Purcell, J. H., & Siegle, D. (2008). Advancement via individual determination: Method selection in conclusions about program effectiveness. *The Journal of Educational Research*, 102, 111-123. doi:10.3200/JOER.102.2.111-124
- Brandon, P. R., & Lawton, B. E., (2013). The development, validation, and potential uses of the Student Interest-in-the-Arts Questionnaire. *Studies in Educational Evaluation*, 39, 90-96. doi.org/10.1016/j.stueduc.2013.01.001
- Castro, R. M. F. (2008). *Establishing a peer and mentoring network to support achievement as an intervention for underachieving gifted Latino high school students* (Doctoral dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3354377)
- Cruse, T. (2012). *Relationship between adolescent gifted girls' attitudes and their value-added performance score* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3507855)

- Davie, J. L. (2012). *Giftedness and underachievement: A comparison of student groups* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3513325)
- Englund, J. M. (2009). *Underachievement among white middle- and high-SES gifted males* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3353823)
- Erkman, F., Caner, A., Sart, Z. H., Börkan, B., & Şahan, K. (2010). Influence of perceived teacher acceptance, self-concept, and school attitude on the academic achievement of school-age children in Turkey. *Cross-Cultural Research: The Journal of Comparative Social Science*, 44, 295-309. doi:10.1177/1069397110366670
- Figg, S. D., Rogers, K. B., McCormick, J., & Low, R. (2012). Differentiating low performance of the gifted learner: Achieving, underachieving, and selective consuming students. *Journal of Advanced Academics*, 23, 53-71. doi: 10.1177/1932202X11430000
- Gelman, A. (2010). *Mario math with millennials: The impact of playing the Nintendo DS on student achievement* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3411899)
- Grahek, M. (2007). *Personality and rater leniency: Comparison of broad and narrow measures of conscientiousness and agreeableness achievement* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 1446591)
- Henry, B. B. (2008). *Factors influencing student achievement in advanced placement and honors courses* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3324980)

- Imasa, L. A. (2012). *An examination of the role of school climate in adolescent test anxiety* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3499684)
- King, L. D. (2011). *An examination of the influence of school discipline, Race/ethnicity, and gender on students' attitudes toward school* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3469144)
- Kirk, C. M., Lewis, R. K., Scott, A., Wren, D., Nilsen, C., & Colvin, D. Q. (2012). Exploring the educational aspirations–expectations gap in eighth grade students: Implications for educational interventions and school reform. *Educational Studies, 38*, 507-519. doi: 10.1080/03055698.2011.643114
- Kusterer, K. D. (2009). *Impact of parenting styles on academic achievement: Parenting styles, parental involvement, personality factors and peer orientation school* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3356956)
- Lang, J. M., Waterman, J., & Baker, B. L. (2009). Computeen: A randomized trial of a preventive computer and psychosocial skills curriculum for at-risk adolescents. *The Journal of Primary Prevention, 30*, 587-603.
- Mason, B. (2008). *Passing CAHSEE but failing math and English: What best predicts high school classroom achievement?* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3341897)
- Matthews, M. S., & McBee, M. T. (2007). School factors and the underachievement of gifted students in a talent search summer program. *Gifted Child Quarterly, 51*, 167-18. doi:10.1177/0016986207299473

- Mombourquette, C. P. (2007). *A study of the relationship between the type of parent involvement and high school student engagement, academic achievement, attendance, and attitude toward school* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3258724)
- Niederdeppe, N. N. (2009). *The impact of a theatre arts intervention on underachieving gifted Latinos* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3351755)
- Olivarez, S. S. (2004). *Factors that predict underachievement in Hispanic gifted students* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 1419158)
- Ragland, D. D. (2008). *Attitude variables affecting student achievement: An assessment of the relationship between student attitudes and achievement among Hispanic high school students* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3327317)
- Rieger, M. M. (2009). *Readiness for college: Factors affecting minority high school students in two large Washington school districts* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3382122)
- Robertson, J. (2011). *Advanced learner perceptions of psychological well-being and school satisfaction in two educational settings* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3438555)
- Rubenstein, L. D. (2011). *Project atlas: Empowering academically underachieving high potential students* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3468067)

- Rubenstein, L. D., Siegle, D. Reis, S. A., McCoach, D. B., & Burton, M. G. (2012). A Complex quest: The development and research of underachievement interventions for gifted students. *Psychology in the Schools, 49*, 678-694. doi: 10.1002/pits.21620
- Simmons, L.A. (2011). Academic performance differences among male and female African American students: An urban high school study (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3439868)
- Suldo, S. M., McMahan, M. M., Chappel, A.M., & Bateman, L. P. (2013). Evaluation of the Teacher-Student Relationship Inventory in American high school students. *Journal of Psychoeducational Assessment* [Electronic version]. doi: 10.1177/0734282913485212
- Suldo, S. M., & Shaffer, E. J. (2008). Looking beyond psychopathology: The dual-factor model of mental health in youth. *School Psychology Review, 37*, 52-68. doi: 10.1177/0734282907303089
- Suldo, S. M., Shaffer, E. J., & Riley, K. N. (2008). A social-cultural-behavioral model of academic predictors of adolescents' life satisfaction. *School Psychology Quarterly, 23*, 56-69. doi: 10.1037/1045-3830.23.1.56
- Suldo, S. M., Shaffer, E. J., & Shaunessy, E. (2008). An independent investigation of the validity of the *School Attitude Assessment Survey-Revised*. *Journal of Psychoeducational Assessment, 26*, 69-82. doi: 10.1177/0734282907303089
- Taheri, A. (2011). An investigation of the effective factors on students' motivational beliefs: The case of Iranian students. *Europe's Journal of Psychology, 7*, 62-80.
- Tenney, J. K. (2011). *Factors affecting the happiness of urban elementary school students: An exploratory study* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3480846)

- Van Eman, L. M. (2009). *Academic adjustment of gifted fifth, sixth, and seventh grade children placed in accelerated math courses* (Doctoral Dissertation). Available from ProQuest Dissertations and Theses database. (UMI No. 3358734)
- Winton, J. B. (2013). *Reversing underachievement among gifted secondary students*. (Doctoral dissertation). Retrieved from <http://hdl.handle.net/10355/37824>

Appendix B

Table B1

Descriptive Statistics for the School Attitude Assessment Survey-Revised (SAAS-R) Items by Pre-IB and IB Groups for Study 1

Factor	Pre-IB					IB					Effect Size
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Sk</i>	<i>K</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Sk</i>	<i>K</i>	
Academic Self-Perceptions											
2. I am intelligent.	161	5.81	1.07	-1.27	1.87	151	5.91	0.99	-1.41	3.68	-0.10
3. I can learn new ideas quickly in school.	161	5.58	1.05	-1.07	2.66	151	5.75	1.00	-1.29	2.85	-0.17
5. I am smart in school.	160	5.59	1.24	-0.99	0.74	150	5.57	1.14	-1.11	1.45	0.02
11. I am good at learning new things in school.	161	5.47	1.10	-0.74	0.45	151	5.62	0.99	-0.88	0.81	-0.14
13. School is easy for me.	161	4.11	1.63	-0.19	-0.85	151	3.87	1.57	0.02	-0.96	0.15
20. I can grasp complex concepts in school.	161	5.24	1.32	-0.86	0.46	149	5.55	0.98	-0.90	1.63	-0.27
22. I am capable of getting straight A's.	161	5.57	1.65	-1.18	0.62	151	5.35	1.71	-0.93	-0.14	0.13
Attitudes Toward Teachers											
1. My classes are interesting.	161	4.87	1.34	-0.89	0.30	151	5.20	1.29	-1.24	1.42	-0.25
9. I relate well to my teachers.	161	4.53	1.41	-0.45	-0.32	151	4.74	1.34	-0.48	0.26	-0.15
14. I like my teachers.	161	4.84	1.44	-0.54	-0.46	151	5.26	1.24	-0.91	0.92	-0.31
16. My teachers make learning interesting.	160	4.27	1.37	-0.27	-0.19	151	4.70	1.31	-0.43	0.15	-0.32
17. My teachers care about me.	161	4.67	1.41	-0.67	0.38	151	5.07	1.33	-0.60	-0.07	-0.29
31. Most of the teachers at this school are good teachers.	161	5.04	1.50	-1.03	0.56	151	5.06	1.36	-0.78	0.34	-0.01

Factor	Pre-IB					IB					Effect Size
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Sk</i>	<i>K</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Sk</i>	<i>K</i>	
34. I like my classes.	161	4.89	1.32	-0.62	0.33	151	5.11	1.09	-0.53	0.21	-0.18
Attitudes Toward School											
6. I am glad that I go to this school.	160	5.62	1.57	-1.30	1.05	151	5.31	1.59	-1.00	0.22	0.20
7. This is a good school.	161	5.70	1.42	-1.37	1.64	151	5.09	1.61	-0.82	-0.11	0.40
12. This school is a good match for me.	161	5.19	1.68	-0.84	-0.19	151	4.89	1.53	-0.55	-0.49	0.19
19. I like this school.	161	5.47	1.51	-1.09	0.62	151	4.87	1.59	-0.60	-0.26	0.39
23. I am proud of this school.	161	5.11	1.53	-0.56	-0.35	151	4.38	1.54	-0.19	-0.17	0.48
Goal Valuation											
15. I want to get good grades in school.	161	6.65	0.82	-2.95	9.09	151	6.65	0.74	-2.82	10.82	0.00
18. Doing well in school is important for my future career goals.	161	6.64	0.69	-2.22	5.62	151	6.54	0.90	-2.52	7.11	0.12
21. Doing well in school is one of my goals.	160	6.43	1.01	-2.46	6.84	150	6.37	0.97	-1.79	3.57	0.06
25. It's important to get good grades in school.	160	6.56	0.85	-2.30	6.20	150	6.44	0.81	-1.59	2.61	0.15
28. I want to do my best in school.	161	6.40	0.93	-1.66	2.27	151	6.42	0.91	-1.71	2.60	-0.02
29. It is important for me to do well in school.	160	6.50	0.85	-1.72	2.34	151	6.48	0.77	-1.70	3.22	0.02

Factor	Pre-IB					IB					Effect Size
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Sk</i>	<i>K</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Sk</i>	<i>K</i>	
Motivation/Self-Regulation											
4. I check my assignments before I turn them in.	160	4.58	1.84	-0.38	-0.98	150	4.33	1.61	-0.32	-0.71	0.14
8. I work hard at school.	161	5.73	1.24	-0.92	0.16	151	5.60	1.41	-1.32	1.55	0.10
10. I am self-motivated to do my schoolwork.	161	5.16	1.67	-0.80	-0.26	151	5.23	1.68	-0.87	-0.10	-0.04
24. I complete my schoolwork regularly.	161	6.14	1.05	-1.33	1.42	151	5.74	1.45	-1.43	1.64	0.32
26. I am organized about my schoolwork.	161	5.17	1.66	-0.75	-0.33	151	5.11	1.72	-0.71	-0.52	0.04
27. I use a variety of strategies to learn new material.	161	4.57	1.58	-0.33	-0.49	151	4.56	1.58	-0.32	-0.61	0.01
30. I spend a lot of time on my schoolwork.	160	5.57	1.48	-0.95	0.23	150	5.73	1.40	-1.25	1.35	-0.11
32. I am a responsible student.	160	5.63	1.22	-1.06	0.74	151	5.70	1.25	-1.57	3.16	-0.06
33. I put a lot of effort into my schoolwork.	161	5.53	1.38	-0.84	0.04	151	5.66	1.19	-1.18	1.74	-0.10
35. I concentrate on my schoolwork.	161	5.30	1.32	-0.55	-0.56	151	5.21	1.33	-0.61	0.05	0.07

Note. Response scale ranged from 1 (*strongly disagree*) to 7 (*strongly agree*). *Sk.* = skewness. *K* = kurtosis. *IB* = International Baccalaureate. Positive effect sizes indicate that pre-IB group means were greater than IB group means.

Appendix B

Table B2

Descriptive Statistics for the School Attitude Assessment Survey-Revised (SAAS-R) Items by Pre-IB and IB Groups for Study 2

Factor	Pre-IB					IB					Effect Size	
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Sk</i>	<i>K</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Sk</i>	<i>K</i>		
Academic Self-Perceptions												
2. I am intelligent.	589	5.81	1.08	-1.32	2.49	560	5.88	0.99	-1.46	3.47	-0.07	
3. I can learn new ideas quickly in school.	589	5.47	1.18	-0.92	0.89	560	5.57	1.12	-1.14	1.80	-0.09	
5. I am smart in school.	588	5.52	1.18	-0.99	0.88	556	5.65	1.07	-1.06	1.51	-0.12	
11. I am good at learning new things in school.	587	5.43	1.24	-0.92	0.74	559	5.55	1.11	-1.08	1.80	-0.10	
13. School is easy for me.	588	3.94	1.65	-0.11	-0.84	559	4.14	1.63	-0.18	-0.77	-0.12	
20. I can grasp complex concepts in school.	586	5.23	1.29	-0.72	0.23	560	5.28	1.28	-1.01	1.13	-0.04	
22. I am capable of getting straight A's.	586	5.50	1.67	-1.00	0.02	560	5.32	1.73	-0.98	0.08	0.11	
Attitudes Toward Teachers												
1. My IB classes are interesting.	589	4.94	1.54	-0.86	0.192	560	4.94	1.47	-1.06	0.62	0.00	
9. I relate well to my IB teachers.	588	4.76	1.47	-0.56	-0.13	559	4.77	1.51	-0.61	-0.14	-0.01	

Factor	Pre-IB					IB					Effect Size
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Sk</i>	<i>K</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Sk</i>	<i>K</i>	
14. I like my IB teachers.	587	5.14	1.42	-0.75	0.09	560	5.20	1.41	-0.89	0.52	-0.04
16. My IB teachers make learning interesting.	589	4.79	1.44	-0.740	0.33	560	4.76	1.41	-0.67	0.20	0.02
17. My IB teachers care about me.	588	5.06	1.43	-0.73	0.35	560	5.06	1.39	-0.81	0.45	0.00
31. Most of the IB teachers at this school are good teachers.	587	5.34	1.44	-1.07	0.80	557	5.33	1.44	-1.06	0.75	0.01
34. I like my IB classes.	587	5.10	1.54	-0.92	0.43	560	4.93	1.52	-0.80	0.18	0.11
Attitudes Toward School											
6. I am glad that I go to this school.	588	5.60	1.47	-1.13	0.78	560	5.06	1.73	-0.75	-0.30	0.34
7. This is a good school.	589	5.40	1.56	-3.98	0.31	559	4.99	1.63	-0.65	-0.34	0.26
12. This school is a good match for me.	588	5.22	1.50	-3.78	0.13	557	4.82	1.67	-0.56	-0.44	0.25
19. I like this school.	587	5.35	1.56	-3.94	0.35	556	4.84	1.71	-0.64	-0.41	0.31
23. I am proud of this school.	587	5.07	1.58	-3.69	0.04	555	4.56	1.67	-0.42	-0.53	0.31
Goal Valuation											
15. I want to get good grades in school.	585	6.63	0.82	-3.15	12.71	559	6.59	0.73	-2.12	4.88	0.05

Factor	Pre-IB					IB					Effect Size
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Sk</i>	<i>K</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Sk</i>	<i>K</i>	
18. Doing well in school is important for my future career goals.	587	6.53	0.87	-2.45	7.03	560	6.47	0.92	-2.63	9.51	0.07
21. Doing well in school is one of my goals.	588	6.39	0.96	-2.08	5.22	559	6.21	1.05	-1.62	2.91	0.18
25. It's important to get good grades in school.	587	6.50	0.92	-2.58	8.49	560	6.34	1.06	-2.53	8.12	0.16
28. I want to do my best in school.	588	6.46	0.86	-1.95	4.65	560	6.25	1.12	-2.07	5.05	0.21
29. It is important for me to do well in school.	586	6.52	0.83	-2.33	7.27	560	6.39	0.93	-2.20	6.67	0.15
Motivation/ Self-Regulation											
4. I check my assignments before I turn them in.	589	4.83	1.63	-0.54	-0.56	558	4.46	1.67	-0.38	-0.91	0.22
8. I work hard at school.	587	5.72	1.24	-1.20	1.50	558	5.47	1.34	-1.01	0.83	0.19
10. I am self-motivated to do my schoolwork.	588	5.35	1.50	-0.91	0.09	560	5.29	1.57	-1.02	0.34	0.04
24. I complete my schoolwork regularly.	588	5.92	1.23	-1.34	1.63	559	5.65	1.37	-1.32	1.48	0.21
26. I am organized about my schoolwork.	585	5.19	1.55	-0.77	-0.09	557	4.85	1.68	-0.54	-0.62	0.21
27. I use a variety of strategies to learn new material.	585	4.88	1.46	-0.58	0.08	560	4.64	1.51	-0.43	-0.47	0.16

Factor	Pre-IB					IB					Effect Size
	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Sk</i>	<i>K</i>	<i>n</i>	<i>M</i>	<i>SD</i>	<i>Sk</i>	<i>K</i>	
30. I spend a lot of time on my schoolwork.	586	5.67	1.32	-0.99	0.45	554	5.46	1.48	-0.96	0.22	0.15
32. I am a responsible student.	588	5.66	1.22	-1.03	0.78	559	5.58	1.29	-1.18	1.45	0.06
33. I put a lot of effort into my schoolwork.	588	5.56	1.34	-0.95	0.52	560	5.42	1.37	-0.90	0.44	0.10
35. I concentrate on my schoolwork.	588	5.33	1.32	-0.80	0.39	560	5.12	1.31	-0.66	0.10	0.16

Note. Items within the Attitudes Toward Teachers subscale were modified to include the term *IB*. Response scale ranged from 1 (*strongly disagree*) to 7 (*strongly agree*). *Sk.* = skewness. *K* = kurtosis. *IB* = International Baccalaureate.

Positive effect sizes indicate that pre-IB group means were greater than IB group means.

Appendix C

Table C1

Unstandardized and Standardized Parameter Estimates for the Five-Factor School Attitude Assessment Survey-Revised Using Robust Maximum Likelihood Estimation for the Pre-International Baccalaureate (IB) Group in Study 1 (n = 161)

Item	Loading	SE	Residual	SE	Stand. Loading	SE	Stand. Residual	SE
ASP								
2	1.00 _a	-	1.09	0.15	0.62	0.07	0.61	0.08
3	0.90	0.16	1.12	0.14	0.66	0.05	0.56	0.07
5	1.23	0.09	0.84	0.12	0.77	0.04	0.41	0.06
11	1.03	0.14	0.62	0.10	0.82	0.04	0.33	0.06
13	1.23	0.18	0.72	0.11	0.80	0.04	0.37	0.06
20	1.22	0.16	1.19	0.18	0.68	0.05	0.53	0.07
22	1.07	0.17	0.75	0.20	0.75	0.07	0.44	0.11
ATT								
1	1.00 _a	-	0.41	0.07	0.80	0.04	0.37	0.07
9	1.12	0.18	0.53	0.12	0.72	0.07	0.48	0.10
14	1.33	0.21	0.42	0.09	0.85	0.04	0.28	0.07
16	1.34	0.21	0.43	0.09	0.80	0.05	0.36	0.08
17	1.35	0.20	1.56	0.19	0.64	0.05	0.59	0.06
31	1.23	0.18	0.66	0.11	0.79	0.05	0.38	0.07
34	1.19	0.20	1.90	0.32	0.55	0.08	0.70	0.09
ATS								
6	1.00 _a	-	0.90	0.27	0.80	0.07	0.37	0.10
7	0.90	0.11	0.74	0.17	0.80	0.05	0.37	0.08
12	1.04	0.09	1.11	0.22	0.78	0.05	0.40	0.08
19	1.09	0.11	0.41	0.12	0.91	0.03	0.18	0.06
23	0.94	0.10	0.94	0.12	0.77	0.04	0.40	0.07
Goals								
15	1.00 _a	-	0.36	0.12	0.68	0.10	0.54	0.14
18	0.87	0.24	0.24	0.05	0.71	0.06	0.50	0.08
21	1.33	0.24	0.47	0.12	0.73	0.06	0.46	0.09
25	1.22	0.26	0.24	0.06	0.81	0.04	0.34	0.07
28	1.37	0.24	0.27	0.05	0.83	0.04	0.31	0.07
29	1.34	0.27	0.16	0.04	0.88	0.03	0.23	0.06
MOT/S-R								
4	1.00 _a	-	2.41	0.26	0.53	0.07	0.72	0.07
8	1.04	0.18	0.51	0.08	0.82	0.04	0.33	0.06
10	1.21	0.20	1.39	0.17	0.71	0.05	0.50	0.07
24	0.77	0.14	0.54	0.08	0.72	0.05	0.49	0.07
26	0.96	0.15	1.87	0.24	0.56	0.06	0.68	0.07

Item	Loading	SE	Residual	SE	Stand. Loading	SE	Stand. Residual	SE
27	0.70	0.11	2.02	0.24	0.43	0.08	0.68	0.07
30	0.96	0.18	1.30	0.20	0.63	0.06	0.60	0.08
32	0.82	0.15	0.84	0.18	0.66	0.07	0.57	0.09
33	1.22	0.18	0.48	0.09	0.86	0.03	0.25	0.05
35	1.09	0.15	0.61	0.11	0.81	0.03	0.35	0.05

Note. ASP = Academic Self-Perceptions; ATT = Attitudes Toward Teachers; ATS = Attitudes Toward School; Goals = Goal Valuation; MOT/S-R = Motivation/Self-Regulation; SE = standard error; Stand. = standardized.

Item loading fixed to 1.0 to identify the model.

Table C2

Unstandardized and Standardized Parameter Estimates for the Five-Factor School Attitude Assessment Survey-Revised Using Robust Maximum Likelihood Estimation for the International Baccalaureate (IB) Group in Study 1 (n = 151)

Item	Loading	SE	Residual	SE	Stand. Loading	SE	Stand. Residual	SE
ASP								
2	1.00 _a	-	0.999	0.15	0.63	0.06	0.61	0.08
3	0.97	0.19	1.19	0.18	0.58	0.07	0.66	0.08
5	1.11	0.22	0.73	0.13	0.72	0.06	0.48	0.08
11	1.31	0.19	0.59	0.11	0.81	0.04	0.35	0.07
13	1.05	0.21	1.04	0.15	0.64	0.07	0.60	0.09
20	0.73	0.73	1.50	0.18	0.43	0.09	0.81	0.08
22	1.02	1.11	0.52	0.08	0.75	0.05	0.44	0.07
ATT								
1	1.00 _a	-	0.47	0.09	0.72	0.07	0.48	0.10
9	1.14	0.20	0.35	0.06	0.81	0.06	0.35	0.08
14	1.35	0.17	0.36	0.07	0.85	0.03	0.28	0.06
16	1.14	0.19	0.30	0.05	0.83	0.04	0.32	0.06
17	1.22	0.24	1.68	0.19	0.56	0.06	0.69	0.07
31	1.05	0.19	0.40	0.07	0.76	0.06	0.42	0.09
34	1.49	0.26	1.80	0.07	0.62	0.06	0.62	0.08
ATS								
6	1.00 _a	-	0.75	0.12	0.84	0.03	0.30	0.06
7	1.01	0.07	0.79	0.13	0.83	0.04	0.31	0.06
12	0.97	0.06	1.64	0.10	0.85	0.03	0.28	0.05
19	1.13	0.07	0.25	0.12	0.95	0.02	0.10	0.03
23	0.91	0.09	0.89	0.17	0.79	0.05	0.38	0.08
Goals								
15	1.00 _a	-	0.31	0.07	0.66	0.08	0.57	0.11
18	1.02	0.21	0.56	0.15	0.56	0.09	0.69	0.10
21	1.59	0.27	0.33	0.07	0.80	0.05	0.35	0.08
25	1.21	0.35	0.30	0.07	0.73	0.07	0.46	0.11
28	1.55	0.38	0.25	0.05	0.84	0.04	0.30	0.07
29	1.22	0.31	0.24	0.06	0.77	0.07	0.41	0.11
MOT/S-R								
4	1.00 _a	-	1.84	0.21	0.53	0.06	0.72	0.07
8	1.23	0.19	0.86	0.15	0.75	0.06	0.44	0.09
10	1.57	0.20	1.01	0.15	0.80	0.04	0.36	0.06
24	1.32	0.18	0.84	0.17	0.78	0.06	0.40	0.09
26	1.48	0.17	1.35	0.21	0.73	0.05	0.46	0.08
27	0.87	0.16	1.92	0.22	0.47	0.07	0.78	0.07
30	0.96	0.16	1.26	0.37	0.59	0.08	0.65	0.10
32	1.20	0.17	0.50	0.09	0.82	0.04	0.32	0.06

Item	Loading	SE	Residual	SE	Stand. Loading	SE	Stand. Residual	SE
33	1.05	0.20	0.60	0.12	0.76	0.07	0.43	0.10
35	1.22	0.19	0.68	0.13	0.78	0.05	0.39	0.08

Note. ASP = Academic Self-Perceptions; ATT = Attitudes Toward Teachers; ATS = Attitudes Toward School; Goals = Goal Valuation; MOT/S-R = Motivation/Self-Regulation; SE = standard error; Stand. = standardized.

aItem loading fixed to 1.0 to identify the model.

Table C3

Unstandardized and Standardized Parameter Estimates for the Five-Factor School Attitude Assessment Survey-Revised Using Robust Maximum Likelihood Estimation for the Pre-International Baccalaureate (IB) Group in Study 2 (n = 589)

Item	Loading	SE	Residual	SE	Stand. Loading	SE	Stand. Residual	SE
ASP								
2	1.00 _a	-	0.59	0.06	0.71	0.03	0.50	0.04
3	1.29	0.10	0.43	0.05	0.83	0.02	0.31	0.04
5	1.22	0.06	0.53	0.05	0.79	0.03	0.38	0.04
11	1.36	0.10	0.45	0.05	0.84	0.02	0.30	0.03
13	1.36	0.12	1.05	0.11	0.63	0.03	0.60	0.04
20	1.29	0.10	1.72	0.07	0.76	0.03	0.43	0.04
22	1.15	1.13	2.00	0.14	0.53	0.04	0.72	0.04
ATT								
1	1.00 _a	-	0.99	0.10	0.76	0.03	0.42	0.04
9	0.98	0.06	0.86	0.07	0.78	0.02	0.40	0.03
14	0.99	0.06	0.69	0.06	0.81	0.02	0.34	0.03
16	1.05	0.05	0.55	0.05	0.86	0.02	0.27	0.03
17	0.90	0.07	0.92	0.09	0.74	0.03	0.45	0.04
31	0.91	0.06	0.93	0.08	0.74	0.03	0.45	0.04
34	1.10	0.05	0.71	0.07	0.84	0.02	0.30	0.03
ATS								
6	1.00 _a	-	0.61	0.07	0.85	0.02	0.28	0.04
7	1.08	0.05	0.62	0.08	0.86	0.02	0.26	0.03
12	1.01	0.05	0.67	0.08	0.84	0.02	0.30	0.04
19	1.15	0.05	0.38	0.05	0.92	0.01	0.16	0.02
23	1.04	0.06	0.82	0.10	0.82	0.02	0.33	0.04
Goals								
15	1.00 _a	-	0.37	0.07	0.67	0.04	0.55	0.06
18	1.05	0.14	0.42	0.06	0.66	0.04	0.56	0.06
21	1.34	0.11	0.40	0.06	0.76	0.04	0.43	0.05
25	1.28	0.15	0.35	0.07	0.76	0.04	0.42	0.06
28	1.31	0.13	0.22	0.03	0.84	0.02	0.30	0.04
29	1.29	0.16	0.19	0.03	0.85	0.02	0.27	0.04
MOT/S-R								
4	1.00 _a	-	1.85	0.13	0.55	0.04	0.70	0.04
8	1.10	0.10	0.57	0.06	0.79	0.02	0.37	0.04
10	1.09	0.10	1.10	0.10	0.71	0.03	0.49	0.04
24	0.98	0.10	0.72	0.07	0.72	0.03	0.48	0.05
26	0.95	0.09	1.68	0.13	0.55	0.03	0.70	0.04
27	0.81	0.08	1.60	0.11	0.50	0.04	0.75	0.04
30	0.85	0.09	1.16	0.03	0.58	0.04	0.67	0.04
32	1.07	0.09	0.56	0.06	0.79	0.02	0.38	0.04

Item	Loading	SE	Residual	SE	Stand. Loading	SE	Stand. Residual	SE
33	1.29	0.11	0.45	0.15	0.87	0.02	0.25	0.03
35	1.20	0.10	0.59	0.06	0.81	0.02	0.34	0.03

Note. ASP = Academic Self-Perceptions; ATT = Attitudes Toward Teachers; ATS = Attitudes Toward School; Goals = Goal Valuation; MOT/S-R = Motivation/Self-Regulation; SE = standard error; Stand. = standardized.

aItem loading fixed to 1.0 to identify the model.

Table C4

Unstandardized and Standardized Parameter Estimates for the Five-Factor School Attitude Assessment Survey-Revised Using Robust Maximum Likelihood Estimation for the International Baccalaureate (IB) Group in Study 2 (n = 560)

Item	Loading	SE	Residual	SE	Stand. Loading	SE	Stand. Residual	SE
ASP								
2	1.00 _a	-	0.41	0.04	0.76	0.03	0.42	0.04
3	1.25	0.01	0.36	0.05	0.85	0.02	0.29	0.04
5	1.06	0.08	0.51	0.07	0.74	0.04	0.45	0.06
11	1.25	0.09	0.35	0.04	0.85	0.02	0.28	0.03
13	1.26	0.11	1.75	0.11	0.59	0.03	0.66	0.04
20	1.31	0.10	0.66	0.07	0.78	0.03	0.40	0.04
22	1.19	1.11	0.66	0.07	0.52	0.04	0.73	0.05
ATT								
1	1.00 _a	-	0.93	0.08	0.76	0.02	0.43	0.04
9	1.07	0.07	0.88	0.09	0.79	0.03	0.38	0.04
14	1.06	0.07	0.59	0.05	0.84	0.02	0.30	0.03
16	1.07	0.05	0.55	0.05	0.85	0.02	0.30	0.03
17	0.93	0.06	0.86	0.07	0.74	0.03	0.45	0.04
31	1.02	0.06	0.79	0.07	0.79	0.02	0.38	0.03
34	1.09	0.06	0.86	0.09	0.79	0.03	0.37	0.04
ATS								
6	1.00 _a	-	0.57	0.06	0.90	0.01	0.19	0.02
7	1.89	0.03	0.75	0.09	0.85	0.02	0.28	0.03
12	0.97	0.03	0.52	0.06	0.90	0.01	0.19	0.04
19	1.04	0.03	0.31	0.04	0.95	0.01	0.11	0.01
23	1.84	0.04	1.09	0.10	0.78	0.02	0.39	0.04
Goals								
15	1.00 _a	-	0.28	0.04	0.69	0.04	0.53	0.06
18	1.30	0.15	0.41	0.06	0.71	0.05	0.49	0.07
21	1.70	0.16	0.37	0.06	0.82	0.03	0.34	0.05
25	1.83	0.17	0.27	0.04	0.87	0.02	0.25	0.04
28	1.78	0.16	0.44	0.08	0.81	0.03	0.35	0.05
29	1.61	0.16	0.20	0.04	0.88	0.03	0.23	0.05
MOT/S-R								
4	1.00 _a	-	1.91	0.12	0.56	0.03	0.69	0.04
8	1.20	0.09	0.54	0.05	0.83	0.02	0.30	0.03
10	1.24	0.11	1.14	0.12	0.73	0.03	0.46	0.05
24	1.11	0.09	0.82	0.08	0.75	0.03	0.44	0.04
26	1.19	0.09	1.60	0.13	0.66	0.03	0.57	0.04
27	0.81	0.07	1.71	0.10	0.50	0.04	0.75	0.04
30	1.19	0.10	0.98	0.09	0.75	0.03	0.45	0.04
32	1.14	0.10	0.54	0.05	0.82	0.02	0.32	0.03

Item	Loading	SE	Residual	SE	Stand. Loading	SE	Stand. Residual	SE
33	1.28	0.09	0.46	0.05	0.87	0.02	0.25	0.03
35	1.16	0.08	0.56	0.06	0.82	0.02	0.33	0.03

Note. ASP = Academic Self-Perceptions; ATT = Attitudes Toward Teachers; ATS = Attitudes Toward School; Goals = Goal Valuation; MOT/S-R = Motivation/Self-Regulation; SE = standard error; Stand. = standardized.

aItem loading fixed to 1.0 to identify the model.