2021 ASEE ANNUAL CONFERENCE
Virtual Meeting | July 26–29, 2021 | Pacific Daylight Time
Paper ID #34167

The Student Attitudinal Success Inventory III (SASI III): Construct Validity and Measurement Invariance

Dr. Jiaqi Zhang, University of Cincinnati
Dr. P.K. Imbrie, University of Cincinnati

P.K. Imbrie is the Head and Professor of the Department of Engineering Education and a Professor in the Department of Aerospace Engineering and Engineering Mechanics University of Cincinnati. He received his B.S., M.S. and Ph.D. degrees in Aerospace Engineering from Texas A&M University. He is an advocate for research-based approaches to engineering education, curricular reform, and student retention. Imbrie conducts both traditional, as well as educational research in experimental mechanics, piezospectroscopic techniques, epistemologies, assessment, and modeling of student learning, student success, student team effectiveness, and global competencies He helped establish the scholarly foundation for engineering education as an academic discipline through lead authorship of the landmark 2006 JEE special reports "The National Engineering Education Research Colloquies" and "The Research Agenda for the New Discipline of Engineering Education." He has a passion for designing state-of-the-art learning spaces. While at Purdue University, Imbrie co-led the creation of the First-Year Engineering Program's Ideas to Innovation (i2i) Learning Laboratory, a design-oriented facility that engages students in team-based, socially relevant projects. While at Texas A&M University Imbrie co-led the design of a 525,000 square foot state-of-the-art engineering education focused facility; the largest educational building in the state. His expertise in educational pedagogy, student learning, and teaching has impacted thousands of students at the universities for which he has been associated. Imbrie is nationally recognized for his work in active/collaborative learning pedagogies, teaming and student success modeling. His engineering education leadership has produced fundamental changes in the way students are educated around the world.

Imbrie has been a member of ASEE since 2000 and has been actively involved with the Society in various capacities. He has served in multiple leadership roles in the ERM and FPD divisions, including: ERM board of directors (2002-2004), program chair for ERM (2005 and 2009), ERM program chair for Frontiers in Education (FIE) (2004), FIE Steering Committee ERM representative (2003-2009), as well as program chair (2016) and division chair (2016-17) for FPD. He has also served on two ASEE advisory committees.

The Student Attitudinal Success Inventory III (SASI III): Construct Validity and Measurement Invariance

Keywords: Engineering education, Goodness of fit, Measurement invariance, Multiple group confirmatory factor analysis, Validity. The Student Attitudinal Success Inventory III (SASI III): Construct Validity and Measurement Invariance

Abstract

The Student Attitudinal Success Inventory (SASI) has been a useful self-reported instrument designed to quantify students' non-cognitive attributes that predict students' success and persistence in Engineering Education. Developed from its first version, the third version of SASI consists 140 items quantifying 16 latent constructs, namely SASI III. The findings in this study provide evidence for the construct validity of the SASI III by talking time variables into consideration, both occasions (pre-survey vs. post-survey) and cohorts (cohort 2018 vs. cohort 2019). Further analysis of measurement invariance (MI) reveals the general pattern of fixed and free factor loadings in the sixteen-factor models.

Using multiple group confirmatory factor analysis (MG-CFA), a series of models have been tested for invariance by comparing goodness-of-fit indexes. MI analysis by cohorts showed supports of scale invariance, while MI analysis by occasions established residual invariance. By providing reliability and construct validity index for the overall validity and each sub-validity and the established minimum scale invariance, SASI III has demonstrated its validity as a useful instrument ready to be used as a comprehensive inventory or for comparison purposes from a longitudinal perspective.

Introduction

The engineer of the twenty-first century will compete in an increasingly global environment and face an expanding array of challenges in business and society (Vest, 2008). To solve those challenge, the engineering education enterprise should produce graduates who are not only technically proficient but also diverse in terms of background, culture, outlook and approach (Moore, Frazier, et al., 2017; Strayhorn, Long III, Williams, Dorime-Williams, & Tillman-Kelly, 2014).

The retention of engineering students, from admission to graduation, has been an chronic international concern in engineering education (Steenkamp, Nel, & Carroll, 2017). Increasing retention of engineering students can potentially increase the number of engineering graduates. Thus, it is crucial to predict or identify students with propensities to drop out of an engineering program, particularly for first-year undergraduate students (Reid, 2009).

Engineering programs admit students based on influential cognitive factors, such as grade point average (GPA) and standardized test scores (e.g., SAT and ACT scores). These factors are used to predict their academic success in a university setting (Yoon, Imbrie, Lin, & Reid, 2014). However, non-cognitive factors have also been identified as integral in students' retention and academic achievement (Al-Sheeb, Hamouda, & Abdella, 2019; Aryee, 2017; Cromley, Perez, & Kaplan, 2016; García-Ros, Pérez-González, Cavas-Martínez, & Tomás, 2019; Williams, Smiley, Davis, & Lamb, 2018). Non-cognitive factors are defined as unobservable traits and latent skills related to students academic achievement (Yoon et al., 2014).

The Student Attitudinal Success Instrument (SASI; Immekus, Imbrie, & Maller, 2004; Immekus, Maller, Imbrie, Wu, & McDermott, 2005; Reid, 2009; Reid & Imbrie, 2008; Yoon et al., 2014) was developed to quantify non-cognitive characteristics of first-year engineering students before entering colleges or universities. The original SASI consisted of 161 items assessing nine specific non-cognitive constructs: 1). intrinsic motivation, 2). academic self-efficacy, 3). expectancy-value, 4). deep learning approach, 5). surface learning approach, 6). Problem-solving approach, 7). leadership, 8). team vs. individual orientation, and 9). major indecision. Later on, five more non-cognitive constructs were added to the SASI II, including 10). goal orientation, 11). implicit beliefs, 12). intent to persist, 13). social climate, 14). self-worth, and 15). career decision. A validation study of SASI II was conducted with over 3400 students enrolled at a large public Midwestern university in 2007 and 2008 in the United States (Yoon et al., 2014). Using an exploratory factor analysis (EFA) in data from the 2007 cohort, and validating the factor structure with confirmatory factor analysis (CFA) using data from the 2008 cohort, a final updated version of SASI, namely SASI II, consisted of 162 items (out of original 246) assessing 15 underlying constructs. After the first development of SASI in 2004, the instrument has been utilized as an important measure to model student retention in the engineering education (see., Imbrie, Lin, Oladunni, & Reid, 2008; Imbrie, Lin, & Reid, 2007; Imbrie, Lin, Reid, & Malyscheff, 2008, 2010; Lin, 2013; Lin, Imbrie, & Reid, 2010).

Purpose and Research Questions

Due to recent evidence, the SASI II has been reduced to 130 items, and there has been a sixteenth factor added to the SASI II data. This third version of SASI, SASI III, consists 140 items quantifying 16 latent constructs. Therefore, it is necessary to validate the SASI III with modified items, because modifications in factors and items potentially change the original psychometric properties. In addition to differences in cohorts, students take SASI III twice during their first semester. The pre-survey takes place before they enter the college, whereas the post-survey is taken at the end of the semester. There are two time-related grouping variables: cohorts (2018 vs. 2019) and occasions (pre-survey vs. post-survey). Whenever psychometric properties are interpreted, it is critical to examine whether the measurement model across groups (e.g., gender, race, time differences) tests the hypothesis that similar interpretation can be derived from the data (Milfont & Fischer, 2010; Van de Schoot, Lugtig, & Hox, 2012). Unfortunately, previous studies fail to consider grouping variables when validating SASI or SASI II.

Thus, this study aims to 1). validate the SASI III with a new set of constructs and items, and 2). examine whether the inventory performance is the same across cohorts and occasions. To accomplish these goals, the following research questions are proposed:

RQ1: What level of reliability for each construct in the SASI III overall, over cohorts and occasions?

RQ2: What is the evidence of construct validity of the SASI III, overall, over cohorts and occasions?

To answer research question one, the internal consistency reliability analysis is conducted for each construct. The index Cronbach's α is reported for the overall data set, each cohort, each occasion, and both cohort and occasion. To answer research question two, a confirmatory factor analysis (CFA) and multiple groups CFA (MG-CFA) are conducted to examine the established factor structure of the SASI III. The goal for CFA and MG-CFA in this study were to 1). provide evidence of the construct validity of the updated SASI II as a whole, 2). test a hypothesized factor structure and evaluate whether the same general factor structure of the SASI III is supported in cohorts and occasions. To accomplish those goals, the analysis specified the proposed SASI III structure model and evaluated multiple model fit indexes, including CFI, TLI, RMSEA, and SRMR, in addition to the χ^2 test.

Theoretical Background

The Student Attitudinal Success Inventory III

The SASI III is the third variation of SASI, and it is an internet-based inventory consisting of 140 items quantifying 16 latent non-cognitive constructs: 1). Academic motivation (AMO), 2). persistence (PST), 3). mastery learning goal orientation (MLG), 4). personal achievement goal orientation (PAG), 5). deep learning approach (DLA), 6). surface learning approach (SLA), 7). problem-solving approach (PSA), 8). implicit beliefs about intelligence and person as a whole (IMB), 9). self-worth in competition (SWC), 10). self-worth in other's approach (SWO), 11). social engagement (SCE), 12). teamwork (TWK), 13). decision making in college major (DMC), 14). fit with major/career (FIT), 15). occupational confidence (OCC), and 16). curiosity and exploration (CEI). The second column in Table 1 shows the number of items in each construct, ranging from 3 in SWO to 24 in AMO.

The options on the SASI III are based on a 6-point Likert scales with the following responses: 1 = Strongly Disagree, 2 = Moderately Disagree, 3 = Disagree Slightly Less than Agree, 4 = Agree Slightly More than Disagree, 5 = Moderately Agree, and 6 = Strongly Agree. All items in SASI III are worded in the same direction, and the response can be interpreted in the same way: a higher value indicates a higher level in the latent construct.

Validity

The validity of an instrument is defined as the degree to which an instrument measures what it should measure (Beavers, Lounsbury, Richards, & Huck, 2013).
Validating an inventory, the construct validity is the foundation, which refers to the degree to which the instrument measures a particular construct. The construct validity can be assessed with the factor analysis (FA).

Factor Analysis

Factor analysis (FA) is a data reduction process assuming many observable variables can be reduced to fewer unobservable (latent) factors that share a common variance (Bartholomew, 1980; Beaujean, 2014). The latent factors are not directly measured but are meaningful and essentially hypothetical constructs that represent observable variables.

Let p denote the number of observed variables $(X_1,...,X_p)$ and m denote the number of unobserved factors $(F_1,...,F_m)$, and $p > m$. X_j is the variable represented in latent factors. Hence, the factor analysis assumes that there are m underlying factors whereby each observed variable is a linear combination of those factors with a residual variance.

$$\begin{aligned} X_{11} &= \lambda_{11}F_1 + \lambda_{12}F_2 + \cdots + \lambda_{1m}F_m + \varepsilon_1 \\ X_{21} &= \lambda_{21}F_1 + \lambda_{22}F_2 + \cdots + \lambda_{2m}F_m + \varepsilon_2 \\ &\vdots \\ X_{p1} &= \lambda_{p1}F_1 + \lambda_{p2}F_2 + \cdots + \lambda_{pm}F_m + \varepsilon_p \end{aligned} \tag{1}$$

where $\Lambda = (\lambda_{11},...,\lambda_{pm})$ denotes the $p \times m$ matrix of factor loadings, that λ_{pm} is the factor loading of p^{th} variable on the m^{th} factor, and the ε_p denotes the unique factor for each variable p. The FA model can be written in a compact matrix form as

$$X_p = \Lambda F_m + E_p \tag{2}$$

where X_p is a $p \times 1$ vector of observed variables, Λ is a $p \times m$ matrix of factor loadings, F_m is a $m \times 1$ vector of unobserved factors, and E_p is the $p \times 1$ vector of unique factor of each observed variable.

The factor loadings provide an idea about how much a specific variable contributes to the factor (Beavers et al., 2013). Factor loadings are very similar to weights in multiple regression analyses representing the correlation between observed variables and unobserved factors.

Confirmatory Factor Analysis

Equation 1 and 2 are called the common factor model, which is the foundation of both exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). While EFA is a descriptive and exploratory procedure, the CFA is a confirmatory process that a specific factor model must be specified. Thus, the CFA is to determine a pre-defined factor model's ability to fit an observed data set (Brown, 2015).

Unlike EFA explores the data to determine the number of common factors and the strength of the relationship between observed variables and unobserved factors, the CFA is useful to 1). establish the validity of a common factor model, 2). compare multiple factor models to account for the same data set, 3). test the significance of specific loadings, 4). test relationships between or among factors, 5). test correlation among factors, and 6). assess the convergent and discriminant validity of observed variables (Beaujean, 2014; Brown, 2015).

There are six necessary steps to perform a CFA: 1). define a factor model, 2). collect the data, 3).

obtain the correlation or variance-covariance matrix of the data, 4). fit the model to the data, 5). evaluate model adequacy, and 6). compare with other models (Brown, 2015). One more step is required to modify the factor model based on some modification indexes in some cases. However, this step is beyond the scope of CFA (Beaujean, 2014).

Measurement Invariance and Multiple Group CFA
Educational and psychological research often compare groups on observed variables and unobserved factors (Milfont & Fischer, 2010). Such studies often include a comparison among 1). multiple points in time (e.g., cross-sectional or longitudinal or both), 2). specific groups of individuals (e.g., gender or race differences), and 3). residents of different countries (e.g., cross-cultural). To obtain a meaningful comparison among groups, the common factor model should be stable across groups (Van de Schoot, Schmidt, De Beuckelaer, Lek, & Zondervan-Zwijnenburg, 2015).

Measurement invariances (MI; Jöreskog, 1971; Mellenbergh, 1989) requires the association between observed variables and unobserved factors to not depend on thememberships or occasions. With MI, differences in estimations among groups do not result from distinctive systematic contents or latent constructs' meanings by groups. MI can be conducted by multiple groups confirmatory factor analysis (MG-CFA; Asparouhov & Muthén, 2014; Hirschfeld & Von Brachel, 2014), which tests the pre-defined CFA model across groups at each constraint, simultaneously. There are degrees of invariance, sequentially from the weakest to the strictest form (Putnick & Bornstein, 2016).

Configural invariance: the first and basic step in MI. This step tests whether the latent factor structures are the same across groups, which indicates that the basic organization of the constructs is established.

Metric (weak) invariance: if configural invariance is established, the next step is to test whether the factor loadings are the same across groups, which indicates that items contribute to latent factors to the same or a similar degree.

Scalar (strong) invariance: if metric invariance is established, the next step is to test whether the intercept of the factor model is the same across groups, which indicates the mean differences in latent factors capture all mean differences in the shared variance of the items.

Residual (strict) invariance: if scalar invariance is established, the next step is to test whether the residuals, both measurement error and residuals that attributes to the unique factor, are the same across groups.

Goodness of Fit
Like CFA, the MG-CFA assesses a pre-defined factor structure. The most frequent approach to evaluate the absolute model fit for the CFA model is to use the Chi-square (χ^2) test to determine how the underlying structure of the existing data differs from the proposed CFA model. The hypothesis underlying the χ^2 test is there is no significant difference between the actual data structure and the proposed CFA model. As a result, if the inventory is truly measuring what it supposes to measure. Then the null hypotheses should not be rejected, $p > 0.05$.

Unfortunately, however, the χ^2 test is sensitive to the sample size (Meade, Johnson, & Braddy, 2008). As a result, relative model fit indexes are utilized to determine the comparable fit, including comparative fit index (CFI), Tucker Lewis Index (TLI), the root mean square error approximation (RMSEA), and the standardized root mean residual (SRMR).

Methods

Sample and Procedure

The SASI III was conducted online using a web-based survey program at a large public Midwestern university in 2018 and 2019. Since the SASI III consists of 140 items, it was evenly split into two parts so that students could finish each piece whenever they were available. Students were asked to participate in the SASI III twice, before their entrance of the program (pre-survey) and after one semester of learning in the program (post-survey). Theoretically, the SASI III measures sixteen constructs by 140 items. The first three columns of Table 1 summarize the non-cognitive constructs measured by the SASI III and the number of items in each construct. The number of items is imbalanced across constructs, ranging from 3 in self-worth in other's approach (SWO) to 24 in academic motivation (AMO).

Table 2 provides demographic information about the samples from two cohorts in the sense of gender and race. In 2018, there was a total of 2066 (N_M = 1606 [77.73%], N_F = 460 [22.27%]) students responded to the online survey, within which 1225 (N_M = 961 [78.45%], N_F = 264 [21.55%]) students finished the pre-survey before their entrance to the college, and 841 (N_M = 645 [76.69%], N_F = 196 [23.32%]) finished the post-survey at the end of their first semester. In 2019, there was a total of 2030 (N_M = 1531 [75.42%], N_F = 499 [24.58%]) students responded to the online survey, within which 1293 (N_M = 994 [76.88%], N_F = 299 [23.12%]) students finished the pre-surveybefore their entrance to the college, and 737 (N_M = 537 [72.86%], N_F = 200 [27.14%]) finished the post-survey at the end of their first semester. The ratio between female and male students was about ¼ and was stable, indicating that the samples were equivalent and balanced across cohorts. However, there were less students who responded to the post-survey compared with the pre-survey within each cohort, indicting that the sample were equivalent but slightly unbalanced across occasions.

Data Analysis

The internal consistency reliability analyses were utilized to answer the first research question. The validity could not be examined before a level of reliability was verified.
Using the psych package (Revelle, 2020) in the statistical software R (R Core Team, 2020), the reliability coefficient of internal consistency, Cronbach's α, were calculated for each construct within the SASI III overall, over cohorts, occasions, and cohorts × occasions.

The confirmatory factor analysis (CFA) and multiple group CFA (MG-CFA) were employed to answer the second research question. The CFA examines whether the pre-defined factor structure fits the data as a whole. In contrast, the MG-CFA tests whether the same factor structure fits the data over cohorts, occasions, and cohorts × occasions. Thus, there were four tests: 1). CFA for data as a whole, 2.) measurement invariance (MI) by cohorts, 3). MI by occasions, and 4). MI by cohorts × occasions.

Since the six-level Likert scales were used in the SASI III, the categorical nature violates the normality assumption. The weighted least square mean and variance adjusted (WLSMV) estimators were utilized to estimate the CFA parameters. Unfortunately, however, as the degree of normality violation increases, the error from WLSMV is inflated dramatically (Suh, 2015). As a result, the root mean square error of approximation (RMSEA) and standardized root mean square error (SRMR) are inflated. Thus, in addition to WLSMV, the maximum likelihood estimation (MLE), which is robust with normality violation when estimating error terms (Chou, Bentler, & Satorra, 1991; Li, 2016), was also utilized to estimate the CFA model by ignoring the categorical nature.

Evaluation criteria: the Cronbach's α normally ranges from 0 to 1. As a rules of thumb in psychometrics (Furr, 2017; Raykov & Marcoulides, 2011), the following criteria were employed to evaluate the internal consistency: > 0.90 – Excellent, > 0.80 – Good, > 0.70 – Accept, > 0.60 – Questionable, > 0.50 – Poor, and ≤ 0.50 – Unacceptable.

The fit indexes were used to evaluate the CFA and MI model fit, including Chi-square (χ^2), comparative fit index (CFI), Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA), and the standardized root mean residual (SRMR). The recommended cut-offs that indicate a good fit are *p*-value > 0.05 for χ^2 test, CFI and TLI ≥ 0.95, RMSEA and SRMR ≤ 0.05 (Kline, 2015; Yu, 2002). However, the Chi-square test is sensitive to the sample size (Bearden, Sharma, & Teel, 1982; VanVoorhis, Morgan, et al., 2007). Given the large sample size in this study, the Chi-square was reported without further interpretation. Again, because the WLSMV estimator inflates the error given the categorical nature of the Likert scale, in this study, looser evaluation criteria were chosen for RMSEA and SRMR, ≤ 0.10 (Hooper, Coughlan, & Mullen, 2008; Xia, 2016). For the robust MLE estimator, RMSEA and SRMR should still be ≤ 0.05, but MLE treats responses as a continuous variable.

Results

This study aims to validate the SASI III that quantifies sixteen non-cognitive constructs with 140 items. Besides, measurement invariance was analyzed to determine if the factor structure meant the same thing to participants across cohorts and occasions.

Descriptive Statistics

Table 2 summarizes demographic information of the participants in gender and ethnicity, along with the percentage. Regarding gender, 1/5 to 1/4 of the participants were female. The sample size showed an imbalance between females and males. Although there were some variances, the ratio was stable regardless of cohorts and occasion. The sample was representative of the rate for students in engineering majors.

Similar results could be driven for ethnicity. The sample was imbalanced across race, but was a representative sample for engineering students.

Internal Consistency

Before conducting CFA, the SASI III's internal consistency was reported in Table 1 for each construct overall, over cohorts, occasions, and cohorts × occasions. Overall, the reliability of the SASI III was good and at least acceptable. The Cronbach's *alpha* ranged from 0.774 for fit with major/career (FIT) to 0.950 for both academic motivation (AMO) and decision making in college major (DMC). Fourteen out of sixteen constructs had a Cronbach's $\alpha > 0.800$, indicating good reliability for the SASI III. Although the reliability coefficient for major/career (FIT) and occupational confidence (OCC) were less than 0.800, they were still within the acceptable range (Cronbach's $\alpha > 0.700$).

The same pattern could be seen for the internal consistency across cohorts, occasions, and cohorts × occasions. This suggests that the SASI III items measure the different constructs that deliver consistent scores regardless of the cohorts and occasions.

Construct Validity

While the reliability indicated how consistent the SASI III was, the CFA was employed to examine the construct validity, whether the collected data confirm the theoretical factor structure. The CFA model was specified in the structure shown in Table 1. The correlation among factors was freely estimated because the oblique rotation was utilized when constructing the inventory (Yoon et al., 2014). Table 3 summarized the CFA results when using data as a whole and separately by both cohorts and occasions.

Overall, the factor loadings were significant when ignoring cohorts and occasions, and all fit indexes were good (χ^2 (9470) = 305882.07, $p < 0.000$, $CFI = 0.96$, $TLI = 0.96$, $RMSEA_{WLSMV}$ = 0.09, $RMSEA_{MLE} = 0.04$, $SRMR_{WLSMV} = 0.08$, $SRMR_{MLE} = 0.07$). That is, the 140 items in the SASI III confirms the 16 hypothesized factor structure. Again, both $RMSEA_{WLSMV}$ and $SRMR_{WLSMV}$ were greater than $RMSEA_{MLE}$ and $SRMR_{WLSMV}$, because WLSMV inflates while MLE approach is robust with the error terms. When using the dataset separately by cohorts and occasions, the model fit indexes showed the same pattern.

When treating samples separately, the factor loadings were significant, and all fit indexes were good. This indicates the data collected from the 2018 pre-survey, 2018 post-survey, 2019 pre-survey, and 2019 post-survey had the same factor structure. But the results provide no evidence that those four subsets are comparable because the factor loadings or variances among factors might be different; even the factor structures are the same. So additional investigation of measurement invariance should be conducted to examine whether those subsets are comparable.

Measurement Invariance

Again, the measurement invariance (MI) is an extension of CFA, examining the invariance of estimated parameters of a set of nested models across multiple groups by employing multiple group CFA (MG-CFA). The measurement invariance is supported or established by examining changes in the goodness of fit indexes when cross-group constraints are imposed on the measurement model.

While CFA tests for the overall fit, the MG-CFA determines if changes in the goodness of fit indexes are meaningful as the constraints increases in MI. In other words, the CFA results in the proceeding section provided evidence of construct validity overall without accounting for group differences. Further investigation is necessary to examine how the group differs in terms of measurement and factor structure. Table 4 summarized fit indexes for MI constraints from the basic configural level to the most strict level.

MG-CFA for cohorts: when considering cohorts only, the strict level of MI was established from the data (χ^2 (19608) = 318574.50, $p < 0.000$, $CFI = 0.96$, $TLI = 0.96$, $RMSEA_{WLSMV} = 0.09$, $RMSEA_{MLE} = 0.05$, $SRMR_{WLSMV} = 0.08$, $SRMR_{MLE} = 0.07$), indicating that the latent factor structures, the factor loadings, the intercepts, and the error variances were the same across cohorts for the 16 factor model.

Because the constraints were added sequentially, once the strict level of MI was supported, the configural, metric, and scalar levels of MI were established.

MG-CFA for occasions: when considering occasions only, the scalar level of MI was established from the data (χ^2 (19608) = 307484.20, $p < 0.000$, $CFI = 0.96$, $TLI = 0.96$, $RMSEA_{WLSMV} = 0.09$, $RMSEA_{MLE} = 0.04$, $SRMR_{WLSMV} = 0.08$, $SRMR_{MLE} = 0.07$), indicating that the latent factor structures, the factor loadings, and the intercepts were the same across cohorts for the 16 factor model. However, the strict level of MI is not supported from the data (χ^2 (19608) = 312838.77, $p < 0.000$, $CFI = 0.89$, $TLI = 0.89$, $RMSEA_{WLSMV} = 0.12$, $RMSEA_{MLE} = 0.10$, $SRMR_{WLSMV} = 0.11$, $SRMR_{MLE} = 0.10$). This meant the error variances were not the same across occasions for the 16 factor model.

Again, because the constraints were added sequentially, once the scalar level of MI was supported, MI's configural and metric levels were established.

MG-CFA for cohorts and occasions: when considering both cohorts and occasions, the results were similar as when only considering occasions. The scalar level of MI was established from the data (χ^2 (39884) = 223109.07, $p < 0.000$, $CFI = 0.96$, $TLI = 0.96$, $RMSEA_{WLSMV} = 0.09$, $RMSEA_{MLE} = 0.05$, $SRMR_{WLSMV} = 0.09$, $SRMR_{MLE} = 0.07$), indicating that the latent factor structures, the factor loadings, and the intercepts were the same across cohorts for the 16 factor model. However, the strict level of MI is not supported from the data (χ^2 (39884) = 332109.07, $p < 0.000$, $CFI = 0.89$, $TLI = 0.89$, $RMSEA_{WLSMV} = 0.11$, $RMSEA_{MLE} = 0.09$, $SRMR_{WLSMV} = 0.12$, $SRMR_{MLE} = 0.10$). This meant the error variances were not the same across occasions for the 16 factor model. Because cohorts had a strict level of MI, and occasions had scalar one. The scalar level of MI considering both cohorts and occasions may be attributed to occasions.

Discussion and Conclusion

This present study aims to examine the psychometric properties of the third version of The Student Attitudinal Success Inventory (SASI III) and provide evidence for construct validity overall, over cohorts, occasions, and cohorts × occasions. Given the importance of the inventory that evaluates academic success and predicts students' persistence, this study contributes to the research by assessing the quality of the inventory with the potential for engineering education and other related fields in higher education.

The SASI III is a 140-item Likert-scale instrument developed to quantify 16 non-cognitive attributes related to academic success and persistence for the first-year engineering students. The literature and theory suggest that sixteen factors contribute to the overall academic success and persistence of first-year engineering students. Previous study and initial validity analyses provide empirical evidence or internal reliability and construct validity for the previous SASI. The internal consistency reliability, construct validity, and measurement invariance of this inventory across two time-related variables were analyzed in the current study.

The internal consistency reliability provides evidence that the SASI III was consistent in the sense of reliability overall, over cohorts and occasions, which also fulfills the pre-requests for examining the construct validity. The omnibus test of construct validity using the confirmatory factor analysis (CFA) provides evidence that the SASI III confirms the theoretical factor structure overall, over cohorts, occasions, and cohorts × occasions. Besides, the multiple group CFA (MG-CFA) that examines measurement invariance (MI) suggests that the strict level of MI is supported across students from different cohorts. Still, only the scalar level of MI is supported when concerning occasions.

Mean Differences in Latent Factors
In practice, once the scalar (strong) invariance is established, researchers are free to compare group mean differences in the latent factors by set the latent factor mean to 0 in one group and estimate the means in other groups. The estimated means represent the differences in latent means compared with the reference group. This study established the scalar level of MI for the SASI III, which provided evidence to compare students by either cohorts or occasions.

Lack of a strict level of MI: while the scalar level of MI is the basic level when conducting further data analyses using the SASI III, the strict level of MI is the golden standard. However, studies on MI of survey scales have shown that the strict level of MI is a challenge to meet (Van de Schoot et al., 2015). The lack of a strict level of MI may cause potential bias that obstructs the comparison of latent factor means.

Full vs. partial invariance: rather than ignore the non-invariance in residual and proceed with tests of mean differences in the latent factors across groups, another common practical approach to accept some degree of violations of MI is to use the partial invariance.

While the full MI examines all possible combinations with groups, the partial invariance investigates measurement invariance with some target combinations of grouping variables. For example, in this study, both cohorts (2018 vs. 2019) and occasions (pre-survey vs. post-survey) were examined, so the full MI simultaneously compares all four combinations of cohorts and occasions: 2018 pre-survey, 2018 post-survey, 2019 pre-survey, and 2019 post-survey.

If the goal was to only compare students in the 2018 pre-survey and 2019 post-survey, the partial invariance might be found between those two combinations rather than all four combinations. However, partial invariance should be used with caution, because standards for partial invariance vary greatly, and there are no empirical studies cited to support different guidelines (Putnick & Bornstein, 2016).

Practical Implications

An inventory such as the SASI III may be pragmatically helpful to quantify non-cognitive attributes in engineering education and other majors under the umbrella of STEM education. Educational professionals may use it as a screening tool for students who are struggling or who may drop out, and policymakers may use the result to decide referring students to educational resources or choosing inventions to address specific concerns.

From a descriptive perspective, students with low scores in academic motivation (AMO), persistence (PST), social engagement (SCE), and decision making in college major (DMA) should be referred to the academic consulting center, those who have low scores in fit with major/career (FIT) could lead to a consulting in the career center. Also, a low score in mastery learning goal orientation (MLG), personal achievement goal orientation (PAG), deep learning approach (DLA), surface learning approach (SLA), and problem-solving approach (PSA) may warrant a referral for tutoring in the major.

The SASI III could be used for intervention development and research for causality-related concerns from an inferential perspective. For example, one of the goals for an academic program is to bridge the gap between high school and university education. With the SASI III results, students who may be at risk or facing challenges in different non-cognitive attributes can be identified or predicted. The admission office can use the results as a reference for newly admitted students to avoid drop out. Academic directors may analyze those results and offer specific recommendations and interventions depending on areas of weakness on the SASI III scale. This will also allow first-year students to seek supports and interventions before their failure in but not limited to engineering majors. Department heads and faculties could also use the results to make programmatic changes as needed. It could also be beneficial to develop professional training programs using the SASI III to assist students.

Finally, from a longitudinal perspective, changes or trends could be detected across (e.g., cohorts) or within (e.g., occasions) generations using the SASI III. For example, success for college students in older generations depends more on individual-related attributes, such as implicit beliefs about intelligence and person as a whole (IMB), self-worth in competition (SWC), and occupational confidence (OCC). On the other hand, younger generation college students may rely more on group-related attributes, like self-worth in other's approach (SCE) and teamwork (TWK).

Limitations and Future Directions

There were several limitations in this study that may impact the results and potential utility of this study.

Firstly, psychometric properties may influence the construct validity through reliability. The reported Cronbach's α coefficients of the SASI III were higher, but the number of items was small on some of the scales (e.g., self-worth in other's approach (SCE), fit with major/career (FIT), social engagement (SCE)). In contrast, for some other scales, there are more items (e.g., academic motivation (AMO), problem-solving approach (PSA), and surface learning approach (SLA)). The imbalance in the number of items in each scale could influence construct validity.

Secondly, this study only considers time as grouping variables (e.g., cohorts and occasions). It is also of interest to examine how the SASI III performs across other nominal variables. For example, if the ultimate goal is to study gender differences in persistence for the first-year engineering students, the multiple group CFA should be conducted over gender. Meanwhile, race or ethnicity should be used when cultural differences are the purpose. Regardless of the goals and what nominal variable is utilized, whether the inventory works equivalent across groups must be investigated before comparing them.

Thirdly, the population may influence reliability and validity since the participants were all from one large public Midwestern university. This limits the generalizability of the results to broader college students. All the sample was in their first year of college, which may affect generalizability beyond the first year of college.

Fourthly, the data included only two cohorts students that are not sufficient to detect generational changes over time. Data from more cohorts should be collected from a longitudinal perspective.

Last but not least, while internal consistency reliability and construct validity focus on the SASI III as a whole test, item response theory (IRT) should be employed to investigate how individual items perform on the SASI III. Based on the IRT results, items with poor performance should be removed without affecting the reliability and validity of the inventory.

References

Al-Sheeb, B. A., Hamouda, A., & Abdella, G. M. (2019). Modeling of student academic achievement in engineering education using cognitive and non-cognitive factors. *Journal of Applied Research in Higher Education*, *11*(2), 178–198.

Aryee, M. (2017). *College students' persistence and degree completion in science, technology, engineering, and mathematics (STEM): The role of non-cognitive attributes of self-efficacy, outcome expectations, and interest* (Unpublished doctoral dissertation). Seton Hall University.

Asparouhov, T., & Muthén, B. (2014). Multiple-group factor analysis alignment. *Structural Equation Modeling: A Multidisciplinary Journal*, *21*(4), 495–508.

Bartholomew, D. J. (1980). Factor analysis for categorical data. *Journal of the Royal Statistical Society: Series B (Methodological)*, *42*(3), 293–312.

Bearden, W. O., Sharma, S., & Teel, J. E. (1982). Sample size effects on chi square and other statistics used in evaluating causal models. *Journal of Marketing Research*, *19*(4), 425–430.

Beaujean, A. A. (2014). *Latent variable modeling using R: A step-by-step guide*. Routledge.

Beavers, A. S., Lounsbury, J. W., Richards, J. K., & Huck, S. W. (2013). Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research, and Evaluation*, *18*(1), 6.

Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. Guilfordpublications.

Chou, C.-P., Bentler, P. M., & Satorra, A. (1991). Scaled test statistics and robust standard errors for non-normal data in covariance structure analysis: a Monte Carlostudy. *British Journal of Mathematical and Statistical Psychology*, *44*(2), 347–357.

Cromley, J. G., Perez, T., & Kaplan, A. (2016). Undergraduate STEM achievement and retention: Cognitive, motivational, and institutional factors and solutions. *PolicyInsights from the Behavioral and Brain Sciences*, *3*(1), 4–11.

Furr, R. M. (2017). *Psychometrics: An Introduction*. Sage Publications.

García-Ros, R., Pérez-González, F., Cavas-Martínez, F., & Tomás, J. M. (2019). Effects ofpre-college variables and first-year engineering students' experiences on academic achievement and retention: a structural model. *International Journal of Technology and Design Education*, *29*(4), 915–928.

Hirschfeld, G., & Von Brachel, R. (2014). Improving Multiple-Group confirmatory factoranalysis in R–A tutorial in measurement invariance with continuous and ordinal indicators. *Practical Assessment, Research, and Evaluation*, *19*(1), 7.

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural Equation Modelling: Guidelines for Determining Model Fit. *Electronic Journal of Business Research Methods*, *6*(1), 53–60.

Imbrie, P. K., Lin, J. J., Oladunni, T., & Reid, K. (2008). Use of a neural network modeland noncognitive measures to predict student matriculation in engineering. In *Annual Conference of American Society for Engineering Education.*

Imbrie, P. K., Lin, J. J., & Reid, K. (2007). Comparison of four methodologies for modeling student retention in engineering. In *Proceeding of the american society for engineering education annual conference & exposition, Honolulu, HI.*

Imbrie, P. K., Lin, J. J., Reid, K., & Malyscheff, A. (2008). Using hybrid data to model student success in engineering with artificial neural networks. In *Proceedings of the Research in Engineering Education Symposium.*

Imbrie, P. K., Lin, J. J., Reid, K., & Malyscheff, A. (2010). Using hybrid data to model student success in engineering with artificial neural networks. In *Proceeding of the american society for engineering education annual conference & exposition, Louisville, KY.*

Immekus, J. C., Imbrie, P., & Maller, S. (2004). The Influence of Pre-College Factors on First-Year Engineering Students' Academic Success And Persistence. In *34th asee/ieee frontiers in education conference, savannah, ga.*

Immekus, J. C., Maller, S. J., Imbrie, P., Wu, N., & McDermott, P. A. (2005). Work in Progress - An Analysis of Students' Academic Success and Persistence Using Pre-college Factors. In *Proceedings frontiers in education 35th annual conference* (pp.S2C–S2C).

Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika*, *36* (2), 109–133.

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. Guilfordpublications.

Li, C.-H. (2016). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior research methods*, *48*(3), 936–949.

Lin, J. J. (2013). *Approaches to modeling student matriculation and retention* (Unpublished doctoral dissertation). Purdue University.

Lin, J. J., Imbrie, P. K., & Reid, K. (2010). Student retention modelling: An evaluation of different methods and their impact on prediction results. In *Proceedings of the Research in Engineering Education Symposium, Palm Cove, Australia.*

Meade, A. W., Johnson, E. C., & Braddy, P. W. (2008). Power and sensitivity of alternative fit indices in tests of measurement invariance. *Journal of applied psychology*, *93*(3), 568.

Mellenbergh, G. J. (1989). Item bias and item response theory. *International journal of educational research*, *13*(2), 127–143.

Milfont, T. L., & Fischer, R. (2010). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of psychological research*, *3* (1), 111–130.

Moore, K., Frazier, R. S., et al. (2017). Engineering education for generation Z. *American Journal of Engineering Education (AJEE)*, *8*(2), 111–126.

Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions andreporting: The state of the art and future directions for psychological research. *Developmental review*, *41*, 71–90.

R Core Team. (2020). R: A Language and Environment for Statistical Computing [Computer software manual]. Vienna, Austria. Retrieved from https://www.R-project.org/

Raykov, T., & Marcoulides, G. A. (2011). *Introduction to psychometric theory*. Routledge.

Reid, K. (2009). *Development of the student attitudinal success instrument: assessment offirst year engineering students including differences by gender* (Unpublished doctoral dissertation). Purdue University.

Reid, K., & Imbrie, P. (2008). Noncognitive characteristics of incoming engineering students compared to incoming engineering technology students: A preliminary examination. In *Proceeding of the american society for engineering education annualconference & exposition.*

Revelle, W. (2020). psych: Procedures for Psychological, Psychometric, and PersonalityResearch [Computer software manual]. Evanston, Illinois. Retrieved from https://CRAN.R-project.org/package=psych (R package version 2.0.7)

Steenkamp, H., Nel, A. L., & Carroll, J. (2017). Retention of engineering students. In *2017ieee global engineering education conference (educon)* (pp. 693–698).

Strayhorn, T. L., Long III, L. L., Williams, M. S., Dorime-Williams, M. L., & Tillman-Kelly, D. L. (2014). Measuring the educational benefits of diversity inengineering education: A multi-institutional survey analysis of women and underrepresented minorities.

Suh, Y. (2015). The performance of maximum likelihood and weighted least square meanand variance adjusted estimators in testing differential item functioning with nonnormal trait distributions. *Structural Equation Modeling: A Multidisciplinary Journal*, *22*(4), 568–580.

Van de Schoot, R., Lugtig, P., & Hox, J. (2012). A checklist for testing measurementinvariance. *European Journal of Developmental Psychology*, *9*(4), 486–492.

Van de Schoot, R., Schmidt, P., De Beuckelaer, A., Lek, K., & Zondervan-Zwijnenburg, M. (2015). Editorial: Measurement invariance. *Frontiers in psychology*, *6*, 1064.

VanVoorhis, C. W., Morgan, B. L., et al. (2007). Understanding power and rules of thumbfor determining sample sizes. *Tutorials in quantitative methods for psychology*, *3* (2), 43–50.
Vest, C. M. (2008). Context and challenge for twenty-first century engineering education. *Journal of Engineering Education*, *97* (3), 235–236.
Williams, R., Smiley, E., Davis, R., & Lamb, T. (2018). The predictability of cognitive andnon-cognitive factors on the retention rate among freshmen college students. *The Journal of Negro Education*, *87* (3), 326–337.
Xia, Y. (2016). *Investigating the chi-square-based model-fit indexes for WLSMV and ULSMV estimators* (Unpublished doctoral dissertation). Florida State University, Tallahassee, FL.
Yoon, S., Imbrie, P., Lin, J., & Reid, K. (2014). Validation of the student attitudinal success inventory ii for engineering students. In *American society for engineering education annual conference & exposition.*
Yu, C.-Y. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes* (Unpublished doctoral dissertation). Universityof California - Los Angeles, Los Angeles, CA.

Table 1
Non-cognitive constructs measured by the SASI III and internal consistency reliability coefficients (Cronbach's α).

ID	Constructs	N_i	Overall	Cohorts		Occasions		2018		2019	
				2018	2019	Pre	Post	Pre	Post	Pre	Post
1	Academic motivation (AMO)	24	0.950	0.950	0.951	0.948	0.943	0.952	0.939	0.944	0.949
2	Persistence (PST)	8	0.894	0.892	0.897	0.882	0.898	0.885	0.891	0.880	0.908
3	Mastery learning goal orientation (MLG)	6	0.888	0.890	0.886	0.869	0.896	0.875	0.895	0.863	0.897
4	Personal achievement goal orientation (PAG)	10	0.936	0.938	0.934	0.928	0.947	0.928	0.949	0.927	0.944
5	Deep learning approach (DLA)	9	0.879	0.882	0.876	0.865	0.892	0.872	0.887	0.858	0.898
6	Surface learning approach (SLA)	13	0.888	0.891	0.884	0.880	0.892	0.888	0.891	0.871	0.894
7	Problem-solving approach (PSA)	14	0.940	0.943	0.937	0.929	0.953	0.933	0.953	0.925	0.953
8	Implicit beliefs about intelligence and person as a whole (IMB)	6	0.886	0.882	0.890	0.870	0.906	0.866	0.900	0.874	0.912
9	Self-worth in competition (SWC)	5	0.885	0.882	0.888	0.879	0.892	0.876	0.887	0.882	0.897
10	Self-worth in other's approach (SWO)	3	0.856	0.846	0.866	0.842	0.877	0.828	0.870	0.856	0.884
11	Social engagement (SCE)	5	0.872	0.870	0.874	0.853	0.871	0.854	0.862	0.851	0.880
12	Teamwork (TWK)	7	0.860	0.871	0.847	0.843	0.877	0.862	0.877	0.822	0.878
13	Decision making in college major (DMC)	10	0.950	0.951	0.948	0.942	0.958	0.943	0.957	0.940	0.958
14	Fit with major/career (FIT)	6	0.774	0.778	0.769	0.766	0.790	0.773	0.788	0.759	0.791
15	Occupational confidence (OCC)	4	0.781	0.782	0.780	0.745	0.839	0.741	0.834	0.749	0.843
16	Curiosity and exploration (CEI)	10	0.883	0.888	0.879	0.875	0.897	0.881	0.898	0.868	0.896
	Total	140									

Note: N_i = number of items in the construct.

Table 2

Demographic information of students on the SASI III

Category		Overall	Cohorts		Occasions		2018		2019	
			2018	2019	Pre	Post	Pre	Post	Pre	Post
Gender										
Female	*n*	959	460	499	563	396	264	196	299	200
	%	23.41	22.27	24.58	22.36	25.10	21.55	23.31	23.12	27.14
Male Race/Ethnicity	*n*	3137	1606	1531	1955	1182	961	645	994	537
	%	76.59	77.73	75.42	77.64	74.90	78.45	76.69	76.88	72.86
American Indian or Alaska Native	*n*	15	14	1	9	6	8	6	1	0
	%	0.37	0.68	0.05	0.36	0.38	0.65	0.71	0.08	0.00
Asian	*n*	319	204	115	193	126	123	81	70	45
	%	7.79	9.87	5.67	7.66	7.98	10.04	9.63	5.41	6.11
Black or African American	*n*	120	74	46	80	40	50	24	30	16
	%	2.93	3.58	2.27	3.18	2.53	4.08	2.85	2.32	2.17
Ethnicity Hispanic/Latino	*n*	108	58	50	73	35	36	22	37	13
	%	2.64	2.81	2.46	2.90	2.22	2.94	2.62	2.86	1.76
Native Hawaiian or other Pacific Islander	*n*	3	0	3	2	1	0	0	2	1
	%	0.07	0.00	0.15	0.08	0.06	0.00	0.00	0.15	0.14
Multiple Race	*n*	64	0	64	41	23	0	0	41	23
	%	1.56	0.00	3.15	1.63	1.46	0.00	0.00	3.17	3.12
Non-resident	*n*	80	0	80	47	33	0	0	47	33
	%	1.95	0.00	3.94	1.87	2.09	0.00	0.00	3.63	4.48
Unknown	*n*	7	0	7	4	3	0	0	4	3
	%	0.17	0.00	0.34	0.16	0.19	0.00	0.00	0.31	0.41
White	*n*	3296	1632	1664	2019	1277	958	674	1061	603
	%	80.47	78.99	81.97	80.18	80.93	78.20	80.14	82.06	81.82
Non-Hispanic	*n*	62	62	0	36	26	36	26	0	0
	%	1.51	3.00	0.00	1.43	1.65	2.94	3.09	0.00	0.00
Non-specified	*n*	22	22	0	14	8	14	8	0	0
	%	0.54	1.06	0.00	0.56	0.51	1.14	0.95	0.00	0.00
Total	*n*	4096	2066	2030	2518	1578	1225	841	1293	737
	%	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Table 3

Confirmatory factor analysis results for construct validity of the SASI III

Fit indices	Overall	2018		2019	
		Pre	Post	Pre	Post
χ^2	305882.07	89142.84	80104.77	80066.36	74950.82
df	9470	9470	9470	9470	9470
p-value	0.00	0.00	0.00	0.00	0.00
CFI	0.96	0.96	0.96	0.95	0.96
TLI	0.96	0.96	0.96	0.95	0.96
RMSEA (WLSMV)	0.09	0.08	0.10	0.08	0.10
RMSEA (MLE)	0.04	0.04	0.05	0.04	0.05
SRMR (WLSMV)	0.08	0.08	0.09	0.08	0.09
SRMR (MLE)	0.07	0.07	0.08	0.07	0.08

Table 4
Measurement invariance results for the SASI III

Category Fit indices	Configural	Metric (Weak)	Scalar (Strong)	Residual (Strict)
By Cohorts				
χ^2	317344.32	319738.48	318574.50	318574.50
df	18940	19064	19608	19608
p-value	0.00	0.00	0.00	0.00
CFI	0.96	0.96	0.96	0.96
TLI	0.96	0.96	0.96	0.96
RMSEA (WLSMV)	0.09	0.09	0.09	0.09
RMSEA (MLE)	0.04	0.04	0.04	0.05
SRMR (WLSMV)	0.08	0.08	0.08	0.08
SRMR (MLE)	0.07	0.07	0.07	0.07
By occasions				
χ^2	301618.31	314184.31	307484.20	312838.77
df	18940	19064	19608	19608
p-value	0.00	0.00	0.00	0.00
CFI	0.96	0.96	0.96	0.89
TLI	0.96	0.96	0.96	0.89
RMSEA (WLSMV)	0.09	0.09	0.09	0.12
RMSEA (MLE)	0.04	0.04	0.04	0.10
SRMR (WLSMV)	0.08	0.08	0.08	0.11
SRMR (MLE)	0.07	0.07	0.07	0.10
By cohorts × occasions				
χ^2	324264.79	341804.77	223109.07	332109.07
df	37880	38252	39884	39884
p-value	0.00	0.00	0.00	0.00
CFI	0.96	0.96	0.96	0.89
TLI	0.96	0.96	0.96	0.89
RMSEA (WLSMV)	0.09	0.09	0.09	0.11
RMSEA (MLE)	0.05	0.05	0.05	0.09
SRMR (WLSMV)	0.09	0.09	0.09	0.12
SRMR (MLE)	0.07	0.07	0.07	0.10