Jurnal Psikologi | E-ISSN: 2302-1098
2024, Vol. 23, No. 2,154-175
https://doi.org/10.14710/jp.23.2.154-175

Testing psychometric property on the Indonesian Academic Self-Efficacy Scale (TIASS)

Kusumasari Kartika Hima Darmayanti[1*], Jesyia Meyrinda[1], Erlina Anggraini[2], Annisa Nurul Utami[3], Ati Suci Dian Martha[4], Lailatul Faizah[3], Dewi Maulina[3], Ali Nina Liche Seniati[3]

[1]Faculty of Psychology, Universitas Islam Negeri Raden Fatah Palembang, South Sumatera, Indonesia

[2]School of Psychology, Northeast Normal University, Changchun, China

[3]Faculty of Psychology, Universitas Indonesia, Depok, West Java, Indonesia

[4]School of Computing, Telkom University, Bandung, West Java, Indonesia

ABSTRACT

Background: First-year university students often face a challenging transition from high school to higher education, struggling to adapt to new academic demands. Previous Studies show that academic self-efficacy—students' belief in their ability to succeed—is a key predictor of their achievement. However, specific instruments to measure this in Indonesia remain underdeveloped.

Purpose: This research aimed to develop and test the psychometric properties of The Indonesian Academic Self-Efficacy Scale (TIASS).

Method: The scale was developed based on Zajacova's theory on academic self-efficacy. Tryout involved 251 first-year college students, both female and male, aged 18-22 (M = 19.56, SD = 2.47). Statistical analysis employed classical test theory and the Rasch model. Reliability was assessed with Stratified Alpha, Cronbach's Alpha, and the Rasch Model, while validity was evaluated through construct and concurrent validity.

Findings: The final 20 items effectively measure academic self-efficacy in first-year undergraduates, demonstrating good reliability (Cronbach's Alpha = .692-.781, Stratified Alpha = .902, Item Separation Reliability = .990, and Person Separation Reliability = .880) and validity (*t-value* range from 9.779 to 36.323 and λ ranged from .533 to .900).

Implication: TIASS is a reliable and valid scale for assessing academic self-efficacy, providing a baseline for designing interventions to boost first-year student's academic confidence.

KEYWORDS
academic self-efficacy; scale; first-year undergraduate students; Indonesian; psychometry

Introduction

The transition to higher education presents significant challenges for students which often induce stress during their first year of undergraduate studies. This period requires adaptation to both a different educational framework and an unfamiliar social environment (Misra & Castillo, 2004). Unlike the structured learning processes characteristic of high school, university students are expected to develop autonomy by independently sourcing references and completing academic assignments

Previous studies about academic self-efficacy in first-year undergraduate students have been conducted by Korgan et al. (2013) and Alegre (2014). Korgan et al. (2013) focused on students within a single academic discipline, limiting the generalizability of their findings to students across diverse majors. Thus, it is not easy to generalize the result to undergraduate students from various majors. Alegre (2014) raised a need for an instrument that is constructed

CONTACT Kusumasari Kartika Hima Darmayanti ✉kusumasari.kartika_uin@radenfatah.ac.id Pangeran Ratu Street, 5 Ulu, Seberang Ulu I, Palembang City, South Sumatera, Indonesia

using more representative samples from different universities to obtain the relationships among motivation processes, academic performance, and other variables. Thus, it is necessary to reconstruct academic self-efficacy measurement for various samples of first-year undergraduate students with various majors, universities, and geographic locations.

Academic self-efficacy is a key factor in supporting the academic success of first-year students. Research has shown that higher levels of academic self-efficacy are associated with improved academic performance (Cobo-Rendón et al., 2020; Hayat et al., 2020). Students with low academic self-efficacy tend to attribute academic failure to personal shortcomings, which may exacerbate negative outcomes (Greco et al., 2022). In contrast, students with high academic self-efficacy are more encouraged to improve their abilities to comprehend lecture materials (Hayat et al., 2020). This positive cycle fosters favorable emotional experiences during learning, further promoting academic achievement and performance (Greco et al., 2022; Hayat et al., 2020).

The concept of academic self-efficacy is grounded in Bandura's social cognitive theory theory (self-efficacy). Bandura (1994) defined self-efficacy as a person's belief in their ability to achieve a predetermined performance that can affect various events in their daily life. Furthermore, Bandura (2006) identified self-efficacy in more specific contexts, including exercise regulation, problem-solving, child development, teaching, and the promotion of skills in reading and mathematics.

Van Rooij et al. (2018) suggested that general self-efficacy is not an appropriate construct to capture self-efficacy within the academic domain. Instead, academic self-efficacy proves to be a more suitable construct, as it is specifically tailored to represent self-efficacy in educational contexts, which general self-efficacy cannot fully encapsulate. Academic self-efficacy is particularly critical for first-year undergraduate students as they navigate the challenges of adapting to the university environment. In this study, academic self-efficacy refers to students' confidence in their ability to carry out such academic tasks as preparing for exams and writing term papers (Zajacova et al., 2005)

A strong sense of academic self-efficacy, which is students' belief in their ability to succeed academically and fully engage in the learning process (Bandura, 1977), proves beneficial for first-year undergraduate students (Grøtan et al., 2019; Hawe et al., 2017; Hayat et al., 2020; Morales-Rodríguez & Pérez-Mármol, 2019). Students with high academic self-efficacy are more likely to interact effectively, perform well both inside and outside the classroom, and balance their priorities such as work, family, and academic responsibilities (Zajacova et al., 2005). Successfully managing these aspects builds confidence in their academic abilities, which, in turn, fosters greater academic success. Based on this explanation, this study refers to the construct of academic self-efficacy proposed by Zajacova et al. (2005), which consists of four dimensions.

According to Zajacova et al. (2005), academic self-efficacy consists of four dimensions, (1) confidence in interacting with the campus environment; including confidence in establishing friendships on campus, interacting with professors and campus staff, getting academic assistance and information, participating in class discussions (i.e., asking and answering), and understanding campus regulations; (2) confidence to perform outside of class; including self-confidence to carry out learning activities, read lecture references, do research, complete lecture assignments on time, prepare for exams, improve writing skills, and understand lecture references; (3) confidence to perform in class; including confidence in succeeding in exams, completing more exams in one week, getting the expected scores, and succeeding in the most difficult classes; (4) confidence to manage work, family, and university affairs; including self-confidence to manage campus and work time, manage time efficiently, manage time to interact with family, and manage time for studying.

Previous measurements have been developed to measure academic self-efficacy. However, research on developing measurements for academic self-efficacy in first-year undergraduate

students still needs to be developed (see Table A1, Appendix A). For example, Van Zyl et al. (2022) developed the General Academic Self-Efficacy (GASE) aimed at students in general. In addition, Dominguez-Lara et al. (2023) also developed the Academic Situation Specific Perceived Self-Efficacy Scale (ASSPSES), which is aimed at university students in Latin America. Furthermore, Greco et al. (2022) constructed the New Academic Self-Efficacy (NASE) aimed at university students in general. In addition, two measurement tools were adapted in the Indonesian version, i.e., the College Academic Self-Efficacy Scale (CASES) and The Academic Self-Efficacy Scale (TASES).

The College Academic Self-Efficacy Scale (CASES) by Owen and Froman (1988) was a measurement tool for academic self-efficacy. Owen and Froman (1988) used Bandura (1977) on self-efficacy to explain academic self-efficacy. In measuring academic self-efficacy, CASES has 33 items. Owen and Froman (1988) added that it is important to modify the academic self-efficacy measurement tools that they made because CASES has deficiencies in measuring academic self-efficacy in students from certain majors, such as art and music. CASES consists of several dimensions, such as (1) confidence in overt behavior on campus in lecture situations, (2) confidence in cognitive operations, such as listening carefully to difficult lecture topics, and (3) confidence in technical capabilities (Owen & Froman, 1988). Ifdil et al. (2019) adapted this measurement in the Indonesian context and found that the Indonesian version of the CASES has good reliability with $\alpha = .931$ and discrimination index $D > .30$.

Sagone and Caroli (2014) developed The Academic Self-Efficacy Scale (TASES), which measures academic self-efficacy. The basic theory of TASES refers to Bandura's self-efficacy (Bandura, 1977). Then, Schunk and Pajares (2002) presented the concept of self-efficacy in academics. TASES consists of four dimensions, i.e., self-engagement, self-oriented decision-making, others-oriented problem-solving, and interpersonal climate. Then, Darmayanti et al. (2021) tested the psychometric property of adapted TASES in Indonesian culture, and it was found to have good reliability and higher validity.

The General Self-Efficacy Scale (GASE) is a 5-item and 5-point Likert scale, a unidimensional construct. Without focusing on first-year students explicitly, this measure is utilized for undergraduate students in the USA and the Netherlands (van Zyl et al., 2022). A 4-point Likert scale and nine items comprise the unidimensional Academic Situation-Specific Perceived Self-Efficacy Scale (ASSPSES, Dominguez-Lara et al., 2023). Although ASSPSES satisfies the psychometric procedures, its application is limited to university students in Latin America. The 44 items on the New Academic Self-Efficacy (NASE) measure have a 5-point Likert scale; 37 items are intended for students currently enrolled in classes, and seven are for those finishing theses. Although the NASE satisfies psychometric procedures, it is not intended for first-year students (Greco et al., 2022).

While instruments exist to measure academic self-efficacy within Indonesian cultural settings, a gap still needs to be in developing a more robust measurement tailored to the Indonesian context. This arises from ensuring that such instruments meet stringent psychometric standards. While both existing instruments (CASES in the Indonesian version and TASES with the Indonesian version) demonstrated good reliability, their validity coefficients need to be more adequate. On the other hand, although both CASES and TASES have been adapted in Indonesia, both measure academic self-efficacy in students without specifying whether undergraduate or graduate students whether they are first-year university students or later. Therefore, there is a need for a new measurement of academic self-efficacy in undergraduate students for first-year university students. This research aimed to develop The Indonesian Academic Self-Efficacy Scale (TIASS) based on Zajacova et al. (2005)'s theory on academic self-efficacy. In addition, this development should satisfy psychometric procedures for scientific method purposes.

Method

Development of the TIASS

For developing TIASS, test development has five stages (Cohen et al., 2013), which are (1) test conceptualization, (2) test construction, (3) test tryout, (4) item analysis, and (5) test revision (Cohen et al., 2013). On test conceptualization, we conceived the idea of developing TIASS for first-year undergraduate students to measure academic self-efficacy. Then, we construct items and score these items. TIASS was established based on the academic self-efficacy theory developed by Zajacova et al. (2005). Academic self-efficacy is multidimensional, and its dimensions are (1) interaction at university, (2) performance out of class, (3) performance in class, and (4) managing work, family, and university (Zajacova et al., 2005). In addition, the TIASS scoring technique used a 4-point Likert scale (1 = very unsuitable to 4 = very suitable; very unsuitable = *sangat tidak sesuai*, unsuitable = *tidak sesuai*, suitable = *sesuai*, and very suitable = *sangat sesuai*).

On the test tryout phase, we tried 64 items for 11 first-year undergraduate students (N = 11; girls = 4, boys = 7) and two experts consisting of a lecturer and researcher at the Department of Educational Psychology, Faculty of Psychology, Universitas Indonesia. Based on the process, the 52 items were good for measuring academic self-efficacy. Finally, this study utilized classical test theory and the Rasch measurement model to analyze items and revise the test. For classical test theory, we applied M*plus* 7.0 (Muthén & Muthén, 2017), and jMetrik 4.1.1 was used for the Rasch measurement model (Meyer, 2014).

Participants and Procedures

In the second phase of the research, 251 participants joined which consisted of 113 first-year undergraduate boys and 138 girls studying in Jakarta, Bogor, Depok, Yogyakarta, Semarang, and Bandung and living in dormitories, boarding houses, and family houses in the age range 18-22 (M = 19.56, SD = 2.47). The effect size in this study refers to Mundfrom et al. (2005) who stated that the sample size of factor analysis is above 100. The sampling method of this study, both online and offline, is a non-probability sampling technique, i.e., accidental sampling, also known as convenience sampling (Cozby & Bates, 2013), for those who voluntarily participate in this study are the participants after fulfilling the informed consent. This technique can be used when the exact population size is unknown and will be discontinued when the sample target has reached the requirement (Cozby & Bates, 2013), with specifications for first-year undergraduate students from public and private universities in Indonesia.

After rechecking the items and revising their readability, researchers prepared for the TIASS trial. Data retrieval was conducted online and offline. Before data collection, each participant was confirmed as a first-year undergraduate student. Afterward, they were asked to sign an informed consent sheet to show their willingness to participate. They were also told to discontinue the test if they were uncomfortable.

Data Analysis

Descriptive Analysis. The descriptive test used in this study is to determine the mean, standard deviation, and normality test. Using Jamovi version 2.3 (The Jamovi Project, 2022), normality testing used coefficient Skewness and Kurtosis ± 2 dan ± 7 (Byrne, 2016; Hair et al., 2019). Here, the mean and standard deviation showed the central tendency and variability (Gravetter et al., 2021).

Reliability Test. In testing reliability, Authors employed Cronbach's alpha coefficient with $\alpha \geq .50$ as an adequate reliability value and a .70–.90 for reliability with high criteria (Hinton et al., 2004). Cronbach et al. (1965) introduced stratified alpha, which measures reliability on multidimensional measures. The stratified alpha method is more appropriate in this study than Cronbach's alpha because Cronbach's alpha tests the reliability of unidimensional measurements (Cronbach, 1951). Therefore, Rae (2007) added that research should conduct Cronbach's alpha

test on each dimension and stratified alpha to calculate the reliability coefficient on all dimensions in multidimensional measurements. Stratified alpha uses Cronbach's alpha coefficient for each dimension, which is then calculated based on the stratified alpha formula (Gignac et al., 2019; Widhiarso & Ravand, 2014). This study calculates stratified alpha using the following formula (Equation 1, Cronbach et al., 1965).

$$\text{Stratified } \alpha = 1 - \frac{\sum \sigma_i^2 (\alpha_i - 1)}{\sigma_x^2}$$

Equation 1

$\sum \sigma_i^2$ is the total variance of each item comprising the dimension, α_i is Cronbach's alpha coefficient of each dimension, and σ_x^2 is the total score variance (Cronbach et al., 1965; Gignac et al., 2019; Rae, 2007). Gignac et al. (2019) stated that the stratified alpha coefficient is estimated at .70 - .77.

In Rasch measurement analysis, the separation index describes reliability. The separation reliability value in the Rasch measurement model reports two things, i.e., item reliability and person reliability. Separation reliability will explain how far the measuring instrument can produce a range of measures on the logit ruler. The value of separation reliability (item or person reliability) will be high if the research sample and the item difficulty level have a wide range and can produce a small measurement error value. The meaning of the outer item means that the item has a level of difficulty from the easiest to the most difficult. Likewise with the research sample, if the sample is broad, it has abilities ranging from the smartest to the least smart (Bond & Fox, 2013).

The basis of consideration in Rasch model analysis is seen from the item fit, difficulty level, Rasch discrimination power, and item information function. Some fit indices provided in Rasch measurement analysis are the item fit is the Mean Square Outfit Value (MnSq), which is accepted as .50 < MNSQ < 1.50, and the *Z*-standard Outfit Value (ZSTD), which is accepted as –2.0 < ZSTD < + 2.0 (Bond & Fox, 2013). If the item on both criteria is met, it is not good and needs to be revised or replaced. Second, item difficulty index < .30 = high difficult, .31 ≤ .70 = moderate, > .70 = easy items. Third, discrimination index ≥ .40 = excellent, .30 ≤ *D* ≤ .39 = good, .20 ≤ *D* ≤ .29 = marginal, and ≤ .190 = poor (Bichi et al., 2019; Bond & Fox, 2013; Clark & Bowles, 2018; Urbina, 2014).

Validity Test. Several approaches to testing validity are testing the construct validity and concurrent validity. The construct validity in this study used Confirmatory Factor Analysis (CFA) by M*plus* version 7.0 (Muthén & Muthén, 2017) and concurrent validity by the Jamovi version 2.3 (The Jamovi Project, 2022). This study used concurrent validity by correlating TIASS with other measurements based on the same construct (Cohen et al., 2013). Concurrent validity was tested by testing the TIASS correlation coefficient with another academic self-efficacy measurement. In this case, this study used an academic self-efficacy measurement (i.e., ASES, Academic Self-Efficacy Scale) developed by Rini et al. (2015), which has a Cronbach alpha reliability of .826 and 13 good corrected items total correlation (*D* ≥ .30) (Crocker & Algina, 2006). The measurement developed by Rini et al. (2015) measures academic self-efficacy as TIASS, with c^2 (62) = 166.395, *p* = .0000, RMSEA = .082, 90% CI = .067–.097, probability RMSEA = .000, CFI = .875, TLI = .842, SRMR = .059, *t*-value = 3.860–19.145, and factor loading = .250–.744. Also, the AVE ranges from .175–.448 (Table 5). Here, the concurrent validity is based on the high correlation coefficient between those measurements (*r* ≥ .50 and *p* < .05; Boos & Stefanski, 2011). In addition, the Average Variance Extracted (AVE) ≥ .50 (Fornell & Larcker, 1981; Hair et al., 2019), which means that if the AVE coefficient is below .50, then there is a higher error in the item to the variance (Hair et al., 2019).

The second is construct validity applied factor analysis, which was used in this study using CFA. Overall, the model was declared fit with the provisions of 90% CI ≤ .05 (Kline, 2023),

probability RMSEA ≥ .05 (Raykov & Marcoulides, 2011), and *t*-value (Est. / *SE*) ≥ 1.96 (Raykov & Marcoulides, 2011), RMSEA (Root Mean Square Error of Approximation) ≤ .06 (Hu & Bentler, 1999), SRMR (Standardized Root Mean Square Residual) ≤ .08 (Hu & Bentler, 1999), CFI (Comparative Fit Index) ≥ .95 (Hu & Bentler, 1999), TLI (Tucker Lewis Index) = 0–1 (Hair et al., 2019), and factor loading (l) ≥ .50 and ideally .70 (i.e., described the level of dimensional accuracy in representing the academic self-efficacy construct) (Hair et al., 2019; Urbina, 2014). Furthermore, researchers also report the chi-square score (χ^2). However, the fit indices are not dependent on the chi-square, as the chi-square was not included in the validation criteria because of its sensitivity to the number of samples (Iacobucci, 2010).

Result and Discussion

Item Analysis

The item analysis process is through the second-order CFA model using the M*plus* 7th version. The 52-item TIASS is divided into four factors (i.e., Factor 1 = Confidence in interacting with the campus environment, Factor 2 = Confidence in performing outside the classroom, Factor 3 = Confidence in performing in the classroom, and Factor 4 = Confidence in managing work, family, and campus affairs) following Zajacova et al. (2005), where each factor has 13 items. The item analysis was then performed in five stages (see Appendix B). Of the 52 items, 20 items were retained based on the factor loading value greater than .50 (Hair et al., 2019).

In the first stage, the 52-item TIASS was tested in the second-order CFA model and 27 items with factor loadings below .50 (Hair et al., 2019) were eliminated. In addition, the second stage showed a second-order CFA model test with three items with factor loadings below .50 (Hair et al., 2019) that were eliminated. The third stage had 22 items, and only one item was eliminated. Meanwhile, the fourth stage had 21 items of TIASS involved in the second-order model test, and the results indicated that one item was eliminated. In the final stage, the 20-item TIASS was valid with an overall factor loading greater than .50 (Hair et al., 2019). Here, psychometric testing was based on the 20-item TIASS.

Descriptive Analysis Results and Reliability Testing

Table 1

Participant Characteristics

Characteristics	*n*	%
Data Collecting		
On-site	231	92%
Online	20	8%
Gender		
Female	138	55%
Male	113	45%
University Background		
Universitas Indonesia	100	40%
UIN Syarif Hidayatullah Jakarta	49	20%
UIN Walisongo Semarang	6	2.40%
Institut Pertanian Bogor	12	5%
Politeknik Negeri Jakarta	25	10%
Politeknik Lembaga Pendidikan Komputer Indonesia Amerika	54	22.60%
Several universities in Bandung	3	1.20%
N/A*	2	0.79%
Residence		
Relatives'/parents' houses	112	45%
Boarding houses	61	24%
Dormitories	53	21%

Islamic boarding schools	15	6%
N/A*	10	4%

*Not Mentioned.

The demographic distribution of participants is presented in Table 1. Here, it can be understood that out of 251 (*N*) first-year undergraduate students, there were 18 females and 2 males from online data collection, and there were 120 girls and 111 boys from offline data collection. Thus, 55% of the participants were girls and 45% were boys. Then, 40% of the participants were first-year undergraduate students from the Universitas Indonesia, 22.6% were first-year undergraduate students from the Politeknik Lembaga Pendidikan Komputer Indonesia Amerika, 20% consisted of first-year undergraduate students from Universitas Islam Negeri Syarif Hidayatullah Jakarta, and the rest came from Universitas Islam Negeri Walisongo Semarang, Institut Pertanian Bogor, Politeknik Negeri Jakarta, and some universities from Bandung City. Lastly, 251 participants lived together with families (*n* = 112, 45%), boarding houses (*n* = 61, 24%), dormitories (*n* = 53, 21%), Islamic boarding school (*n* = 15, 6%), and others (*n* = 10, 4%).

Table 2 presents the results of the reliability test and descriptive analysis. Hair et al. (2019) show that data distribution is normal in the skewness coefficient ± 2 range and kurtosis coefficient ± 7. Therefore, each dimension of TIASS and TIASS has normally distributed data. In more detail, interaction at university indicates (*M* = 2.980, *SD* = .525) with normally distributed data (Skewness = −.085, Kurtosis = −.013). Then, performance out of class indicated (*M* = 3.050, *SD* = .467) with normal data (Skewness = .133, Kurtosis = −.279). On performance in class, this study found (*M* = 2.950, SD = .455) and normality test results (Skewness = .090, Kurtosis = −.314). Managing, work, family, and university obtained results of (*M* = 3.020, *S*.D. = .528) and (Skewness = −.111, Kurtosis = −.238). In addition, TIASS showed (*M* = 2.990, *SD* = .390) with (Skewness = .352, Kurtosis = −.956).

Meanwhile, the reliability test results are shown using stratified alpha and Cronbach's alpha coefficients. Taber (2018) stated that the internal consistency coefficient above .60 is acceptable, while Nunnally and Bernstein (1994) stated that Cronbach's alpha coefficient above .70 is reliable. In this study, the coefficient α = .761, .692, .796, and .784 (on interaction at university, performance out of class, performance in class, and managing work, family, and university). Meanwhile, the calculation of stratified alpha (α_s) in the study is as follows (Equation 2).

$$\alpha_s = 1 - \frac{\{6.890*(1-0.761)\}-\{3.489*(1-0.692)\}-\{10.133*(1-0.796)\}-\{5.234*(1-0.781)\}}{60.865}$$

$\alpha_s = .902$

Equation 2

Based on the above calculations in Equation 2 and details below in Table 2, the stratified alpha coefficient (α_s) is .902. The stratified alpha coefficient is estimated above .70-.77 (Gignac et al., 2019). Also, Nunally (1978) stated that the cut-off for alpha is above .70, and coefficient alpha above .60 is acceptable (Taber, 2018). Therefore, it can be concluded that the 20-item TIASS is reliable.

Table 2

Descriptive Statistic and Reliability Testing Results

Dimensions/Variable	α_s	α	*Mean ± SD*	*Skewness*	*Kurtosis*
Interaction at university (item 1-item 5)	-	.761	2.980 ± .5250	-.085	-.013
Performance out of class (item 6-item 9)	-	.692	3.050 ± .467	.133	-.279
Performance in class (item 10-item 16)	-	.796	2.950 ± .455	.090	-.314

Managing work, family, and university (item 17-item 20)	-	.781	3.020 ± .528	-.111	-.238
The Multidimensional TIASS	**.902**	**-**	**2.990 ± .390**	**.352**	**-.956**

Note. α_s = Stratified Alpha.

Based on Table 3, it can be seen that item bias, according to Bond and Fox (2018), is not the main focus in item selection. However, information about the existence of biased items greatly influences measurement accuracy. An item called bias will impact measurement accuracy (Urbina, 2014). An item is called biased if it is found that individuals with certain characteristics are more favored in answering questions than individuals with other characteristics. In the Rasch measurement model, item bias can be detected with DIF (differential item functioning). Items identified with DIF ($p < .05$) are recommended to be reviewed and, if necessary, revised or replaced. TIASS has a DIF value above $p < .05$, as shown in Table 3.

Table 3.
Psychometric Properties of TIASS at The Item Level (N = 251)

Item Number	Factor loading	h^2 *	Infit MnSq	Outfit MnSq	Difficulty	DIF Contrast across Gender [c, d]
Item 1	.45	.48	1.08	1.07	-.21	.50
Item 2	.40	.27	1.37	1.41	.29	.53
Item 3	.13	.34	.97	.97	.30	.72
Item 4	.98	.45	1.15	1.13	-.21	.32
Item 5	.81	.61	1.37	1.35	-.03	.37
Item 6	.35	.43	.85	.92	-.24	.55
Item 7	.19	.52	.86	.88	-.30	.66
Item 8	.12	.30	.81	.90	-.05	.73
Item 9	.30	.47	.83	.83	-.05	.59
Item 10	2.61	.45	.90	.92	-.47	.11
Item 11	.38	.55	.96	.98	.05	.54
Item 12	.18	.51	.84	.87	.30	.67
Item 13	3.09	.64	.99	.99	.58	.58
Item 14	.53	.36	.83	.84	-.16	.47
Item 15	.60	.35	.85	.86	-.02	.44
Item 16	1.10	.41	.82	.82	.67	.29
Item 17	.96	.30	1.17	1.24	.67	.33
Item 18	.00	.64	1.15	1.12	-.15	.96
Item 19	.05	.75	1.04	1.02	-.33	.82
Item 20	6.20	.42	1.11	1.12	-.61	.51

Note. MnSq = mean square error; DIF = differential item functioning; h^2 = communalities.
* Extraction method: Oblimin rotation with Kaiser normalization; [c] DIF contrast > .5 indicates substantial DIF; [d] DIF contrast across gender = difficulty for males-difficulty for females

Validity Testing Results

Table 4 is a summary of the second-order model for the TIASS. The 20 items are valid in measuring academic self-efficacy with a *t*-value range (9.779–36.323) with a *t*-value ≥ 1.96 (Raykov & Marcoulides, 2011). The second-order model results showed that the model fit with $\chi^2(166) = 379.021$, $p = .0000$, RMSEA = .072, 90% CI = .062– .081, probability RMSEA = .000, CFI = .872, TLI = .853, SRMR = .065. Hu and Bentler (1999) mention that RMSEA is below .06, but Hair et al. (2019) said that the RMSEA coefficient ranges from .03 to .08. Xia and Yang (2019) stated

that RMSEA > .80 is a reasonable fit. Meanwhile, this study indicates that RMSEA = .072. Then, the SRMR in this study was found to be .065 (below .08). Thus, the absolute fit measure has achieved goodness of fit. For incremental fit indices, CFI and TLI are in the range of 0−1, with higher coefficients, the more fit (Hair et al., 2019).

Table 4
Second-order Model of the TIASS (N = 251)

Item	*Factor Loading*	*Est./SE*	*p-value*
Item 1. I am not shy to communicate with lecturers. *Saya tidak malu untuk berkomunikasi dengan dosen.*	.682	15.145	.000
Item 2. I do not hesitate to strike up a conversation with the campus security guard. *Saya tidak ragu untuk memulai perbincangan dengan satpam kampus.*	.533	9.779	.000
Item 3. I do not hesitate to ask the academic department about semester credit units and study plan cards *Saya tidak ragu untuk bertanya pada bagian akademik terkait Satuan Kredit Semester (SKS) dan Kartu Rencana Studi (KRS).*	.619	12.493	.000
Item 4. I believe I can express my opinion during class discussion sessions. *Saya yakin dapat menyampaikan pendapat ketika sesi diskusi di kelas.*	.611	12.228	.000
Item 5. I am not shy to ask questions to the lecturer in class. *Saya tidak malu untuk mengajukan pertanyaan kepada dosen di kelas.*	.697	15.541	.000
Item 6. I can complete the paper before the deadline that has been set. *Saya percaya dapat menyelesaikan paper sebelum deadline yang telah ditetapkan.*	.534	10.178	.000
Item 7. I can study hard so that I can do well on the exam. *Saya yakin dapat belajar dengan giat supaya mampu mengerjakan ujian dengan lancar.*	.593	12.111	.000
Item 8. I am confident that I can make a preparation plan for midterm exam and end-of-semester exam *Saya yakin dapat membuat rencana persiapan untuk menghadapi UTS dan UAS.*	.647	14.276	.000
Item 9. I can improve my reading speed by using certain strategies. *Saya yakin dapat mengasah kecepatan dalam membaca dengan menggunakan strategi tertentu.*	.621	13.501	.000
Item 10. I can pass the midterm exam successfully. *Saya percaya dapat melewati masa Ujian Tengah Semester dengan sukses.*	.580	12.025	.000
Item 11. I am confident that I can answer the end-of-semester exam questions correctly *Saya yakin dapat menjawab soal UAS dengan benar.*	.589	12.270	.000
Item 12. I can do the quizzes continuously for one week.	.605	13.104	.000

Saya yakin dapat mengerjakan kuis yang dilaksanakan *bersusulan dalam satu minggu.*			
Item 13. I am confident that I can cope with a heavy exam load in a week. *Saya yakin dapat menghadapi beban ujian yang padat dalam sepekan.*	.537	10.520	.000
Item 14. I believe in getting the assignment grades I want. *Saya percaya untuk mendapatkan nilai tugas yang sesuai keinginan.*	.663	15.972	.000
Item 15. I am confident that I can get a satisfactory quiz score *Saya yakin bisa memperoleh nilai kuis yang memuaskan.*	.639	14.464	.000
Item 16. I am confident that I can successfully master the most complicated courses. *Saya yakin dapat berhasil menguasai mata kuliah yang paling rumit.*	.590	12.532	.000
Item 17. I do not find it difficult to set aside time to study. *Saya tidak merasa kesulitan untuk menyediakan waktu khusus untuk belajar.*	.547	11.262	.000
Item 18. I am confident that I can manage my time between family affairs and coursework in a balanced manner. *Saya percaya diri dapat mengatur waktu antara urusan keluarga dan tugas kuliah secara seimbang.*	.841	30.957	.000
Item 19. I am confident that I can make time to contact my family members while I am busy studying. *Saya yakin dapat meluangkan waktu untuk menghubungi anggota keluarga di sela-sela kesibukan kuliah.*	.900	36.323	.000
Item 20. I believe I can set aside special time to share with my family. *Saya percaya dapat menyediakan waktu khusus untuk sharing dengan keluarga.*	.525	10.478	.000

The second-order diagram of the TIASS model can be seen in Figure 1. Figure 1 shows that the factor loading of 20 items ranges from .533 to .900. According to Hair et al. (2019), factor loading is at least .50 and ideally .80. Therefore, 20 items are valid in measuring academic self-efficacy.

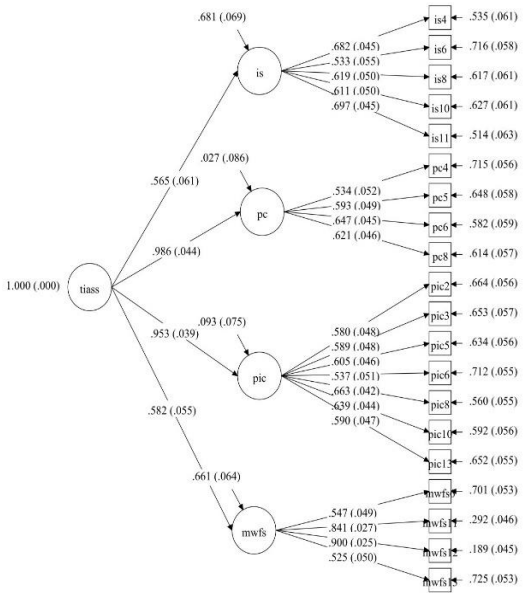

Figure 1. *Second-order Model of TIASS (N = 251)*

Table 5 is the result of concurrent validity by correlating TIASS with ASES (r = .605, p < .001), so TIASS is declared valid with a validity coefficient (r ≥ .50 and p < .05) (Boos & Stefanski, 2011). In detail, interaction at university (r = .355, p < .001) and managing, work, family, and university (r = .412, p < .001) have correlation coefficients with ASES below r = .50, so both dimensions have low validity coefficients. Meanwhile, performance out of class (r = .526, p < .001) and performance in class (r = .572, p < .001) indicate that both dimensions accurately estimate an individual's academic self-efficacy level on the relevant criterion (i.e., ASES). In addition, from Table 5, it can be concluded that only managing, work, family, and university (AVE = .528), which indicates that (AVE) ≥ .50 (Fornell & Larcker, 1981; Hair et al., 2019), is valid. On the other hand, interaction at university (AVE = .395), performance out of class (AVE = .363), and performance in class (AVE = .359) showed poor validity.

Table 5

Correlation between TIASS and ASES and Average Variance Extracted Coefficient

Dimensions /Variable	1	2	3	4	5	6	7	8	9	AVE
IU	-									.395
POC	.435***	-								.363
PIC	.432***	.694***	-							.359
MWFU	.310***	.487***	.514***	-						.528
TIASS	**.703*** **	**.810*** **	**.869*** **	**.725*** **	**-**					**-**
Generality	.293***	.463***	.431***	.356***	.496***	-				.175
Magnitude	.259***	.448***	.487***	.35***	.503***	.633***	-			.348
Strength	.345***	.426***	.535***	.342***	.539***	.480***	.627***	-		.448
ASES	**.355*** **	**.526*** **	**.572*** **	**.412*** **	**.605*** **	**.835*** **	**.877*** **	**.833*** **	**-**	**-**

Note. IU = Interaction at University; POC = Performance Out of Class; PIC = Performance in Class; MWFU = Managing Work, Family, and University.

***p < .001

This study aims to prove that the development of TIASS found that TIASS has a stable unidimensional structure with strong psychometric properties. The results of preliminary

psychometric tests showed that TIASS has good properties regarding the type of testing utilized by CTT analysis and the Rasch model. In addition, the overall score of the summed item scores can indicate high academic self-efficacy.

The psychometric property tests in this study used reliability and validity tests. The reliability test involves Cronbach's Alpha, composite reliability, and the Rasch measurement model. The validity test uses concurrent validity and construct validity. To be more detailed, factor analysis resulted in construct validity, part of the CTT. Furthermore, the Rasch measurement model, which shows reliability and item analysis, is part of modern test theory.

When viewed from the coefficient Cronbach's Alpha, stratified alpha, and Rasch measurement model, TIASS is a reliable measurement tool, as summarized in Table 6. One basic assumption of the Rasch measurement model is unidimensionality. To ensure the test measures the intended objective, assessing unidimensionality is crucial. The PCA of the Rasch residuals was performed to determine the unidimensionality of this study. The raw variance explained by measures is 24.9%, which closely matches the expected variance of 24.7%. The raw variance explained by a person is 5.8%, and the variance explained by items is 19.89%. The results show that the variance explained of 24.9% is higher than the minimum unidimensionality requirement of 20%; this means that the unidimensionality is achieved and the test measures unidimensional constructs.

Table 6

Psychometric Properties of TIASS at The Item Level (N = 251)

Psychometric testing	Coefficient	Suggested cutoff
Average variance extracted	.363 to .528	> .5
Internal consistency (Cronbach's α)	.692 to .781	> .7
Stratified α	.902	> .7
Standard error of measurement	.044	The smaller, the better
Item separation reliability from Rasch	.990	> .7
Item separation index from Rasch	4.420	> 2
Person separation reliability from Rasch	.880	> .7
Person separation index from Rasch	2.810	> 2

From the results of the Rasch measurement model analysis, we obtained information for the reliability index and person separation of .88 and for the person separation value of 2.81. This reliability value is classified as good; this indicates that the variability of student abilities in this study is sufficient. This shows that the ability of academic self-efficacy is well tested, and there are three different groups of students, namely students with low, medium, and high skills. Furthermore, the item reliability and separation index values are .90 and 4.42. Both values indicate that the reliability of the items on the development scale is very good, and the sample of people is large enough to confirm the hierarchy of the difficulty level of the test items (Figure 2).

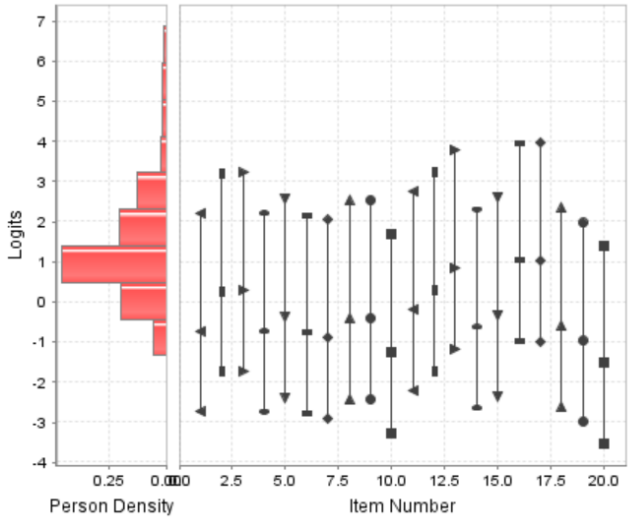

Figure 2. *Items Distribution of TIASS*

Table 7 also shows that the 20-item TIASS validates academic self-efficacy based on a factor loading above .50 (Hair et al., 2019) and a *t*-value above 1.96 (Raykov & Marcoulides, 2011). In addition, when looking at the summary fit indices of the TIASS in Table 7, the second-order model of the TIASS was found to be fit. Thus, the 20 items fulfill the procedures of a valid measurement. In addition, Table 7 also shows the comparison of fit indices of TIASS between 52 items and 20 items.

In the 52-item TIASS, some items still have factor loading coefficients below .50 and *t*-values below 1.96. Based on those considerations, 22 items in the TIASS with factor loading coefficients and *t*-values below .50 (Hair et al., 2019) and 1.96 (Raykov & Marcoulides, 2011) were eliminated. Thus, when looking at Table 7, it can also be found that there is an increase in fit indices on the 20-item TIASS rather than the 52-item TIASS.

Table 7
The Comparison of Fit Indices on CFA Based on Hair et al. (2019)

Fit Indices	Value	Cutt-off Value
Fit Indices 52-item TIASS		
Factor loading	-.070 to .774	≥ .50
t-value (Est. / *SE*)	-1.030 to 21.273	≥ 1.96
RMSEA (Root Mean Square Error of Approximation)	.068	.03 to .08
SRMR (Standardized Root Mean Square Residual)	.082	< .1
CFI (Comparative Fit Index)	.618	0-1
TLI (Tucker Lewis Index)	.601	0-1
Fit Indices 20-item TIASS		
Factor loading	.533 to .900	≥ .50
t-value (Est. / *SE*)	9.779 to 36.323	≥ 1.96
RMSEA (Root Mean Square Error of Approximation)	.072	.03 to .08
SRMR (Standardized Root Mean Square Residual)	.065	< .1
CFI (Comparative Fit Index)	.872	0-1
TLI (Tucker Lewis Index)	.853	0-1

Based on the previous explanation, TIASS is a psychological measurement tool used to measure academic self-efficacy in first-year undergraduate students in Indonesia. This

measurement is declared valid and reliable based on the results of psychometric property tests using the CTT and Rasch measurement models. Previously, academic self-efficacy research and assessment used several tools developed outside Indonesia.

Van Zyl et al. (2022) developed the General Academic Self-Efficacy (GASE), which consists of five items. This measurement is unidimensional (Van Zyl et al., 2022). The validity test used in this study uses construct validity, concurrent validity, and predictive validity (Van Zyl et al., 2022). This measurement was developed for people in Western Europe and the U.S. (Van Zyl et al., 2022). Also, GASE is used for university students, but this measurement is not used specifically for first-year undergraduate students (Van Zyl et al., 2022).

One of the widely adapted academic self-efficacy measurement tools in the Indonesian context is the Collage Academic Self-Efficacy Scale (CASES), developed by Owen and Froman (1988) and adapted by Ifdil et al. (2019). Owen and Froman (1988) developed this measurement for students in general, one of which was used in Zysberg and Schwabsky (2021). This measurement has high internal consistency ($\alpha = .931$) in the Indonesian version (Ifdil et al., 2019).

A measurement of academic self-efficacy in first-year university students was found by Byrne et al. (2014). Byrne et al. (2014) constructed a measurement only for accounting students in this study. Therefore, academic self-efficacy has yet to be measured among first-year undergraduate students in Indonesia. TIASS was developed as an assessment and research related to the academic self-efficacy of first-year undergraduate students in Indonesia. The measurement results using TIASS can also be used as a baseline for providing psychological interventions in dealing with first-year undergraduate students to improve academic achievement and learning motivation.

Conclusion

This study aimed to develop TIASS and test its psychometric properties so that TIASS can meet the criteria of a good measurement and fulfill psychometric procedures. The TIASS construction stage consists of test conceptualization, test construction, test tryout, item analysis, and item revision. After determining 64 items that measure academic self-efficacy in first-year undergraduate students, 52 items were retained during the test tryout stage. At the item analysis and revision stage, psychometric property tests were applied using the CTT and Rash measurement model approaches. From the item analysis and revision stage, 20 TIASS items were valid and reliable. This indicates that the TIASS was found to meet the procedures of psychometrics and can be used for research and as a baseline in conducting psychological interventions.

References

Alegre, A. A. (2014). Academic self-efficacy, self-regulated learning and academic performance in first-year university students. *Propósitos y Representaciones*, *2*(1), 101–120. https://doi.org/10.20511/pyr2014.v2n1.54

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Bandura, A. (1994). Self-efficacy. In *Encyclopedia of human behavior* (4th ed., pp. 71–81). Academic Press.

Bandura, A. (2006). *Self-efficacy beliefs of adolescents*. Information Age Publishing.

Bichi, A. A., Talib, R., Atan, A., & Ibrahim, H. (2019). Validation of a developed university placement test using classical test theory and Rasch measurement. *International Journal of Advanced and Applied Sciences*, *6*(6), 22–29. https://doi.org/10.21833/ijaas.2019.06.004

Bond, T. G., & Fox, C. M. (2013). *Applying the Rasch model*. Psychology Press. https://doi.org/10.4324/9781410614575

Boos, D. D., & Stefanski, L. A. (2011). P -Value precision and reproducibility. *The American Statistician*, *65*(4), 213–221. https://doi.org/10.1198/tas.2011.10129

Byrne, B. M. (2016). *Structural equation modeling with Amos: Basic concepts, applications, and programming*. Routledge.

Byrne, M., Flood, B., & Griffin, J. (2014). Measuring the academic self-efficacy of first-year accounting students. *Accounting Education*, *23*(5), 407–423. https://doi.org/10.1080/09639284.2014.931240

Clark, D. A., & Bowles, R. P. (2018). Model fit and item factor analysis: Overfactoring, underfactoring, and a program to guide interpretation. *Multivariate Behavioral Research*, *53*(4), 544–558. https://doi.org/10.1080/00273171.2018.1461058

Cobo-Rendón, R., Pérez-Villalobos, M. V., Páez-Rovira, D., & Gracia-Leiva, M. (2020). A longitudinal study: Affective wellbeing, psychological wellbeing, self-efficacy and academic performance among first-year undergraduate students. *Scandinavian Journal of Psychology*, *61*(4), 518–526. https://doi.org/10.1111/sjop.12618

Cohen, R. J., Swerdlik, M. E., & Sturman, E. D. (2013). *Psychological testing and assessment: an introduction to tests and measurement* (8th ed.). McGraw-Hill Education.

Cozby, P. C., & Bates, S. C. (2013). *Methods in behavioral research* (12th ed.). McGraw-Hill.

Crocker, L., & Algina, J. (2006). *Introduction to classical and modern test theory*. Cengage Learning.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/10.1007/BF02310555

Cronbach, L. J., Schönemann, P., & McKie, D. (1965). Alpha coefficients for stratified-parallel tests. *Educational and Psychological Measurement*, *25*(2), 291–312. https://doi.org/10.1177/001316446502500201

Darmayanti, K. K. H., Anggraini, E., Winata, E. Y., & Mardianto, M. F. F. (2021). Confirmatory factor analysis of the academic self-efficacy scale: An Indonesian version. *Jurnal Pengukuran Psikologi Dan Pendidikan Indonesia (JP3I)*, *10*(2), 118–132. https://doi.org/10.15408/jp3i.v10i2.19777

Dominguez-Lara, S., Alarcón-Parco, D., Campos-Uscanga, Y., Tamayo-Agudelo, W., Merino-Soto, C., Tumino, M. C., Quinde, J. M., De Moura, G. B., & Baptista, M. das G. de A. (2023). Psychometric properties and measurement invariance of an academic self-efficacy scale in college students from five Latin American countries. *Ciencias Psicológicas*. https://doi.org/10.22235/cp.v17i1.3051

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*, 39–50. https://doi.org/10.1177/002224378101800104

Gignac, G. E., Reynolds, M. R., & Kovacs, K. (2019). Digit span subscale scores may be insufficiently reliable for clinical interpretation: Distinguishing between stratified coefficient alpha and omega hierarchical. *Assessment*, *26*(8), 1554–1563. https://doi.org/10.1177/1073191117748396

Gravetter, F. J., Wallnau, L. B., Forzano, L. A. B., & Witnauer, J. E. (2021). *Essentials of statistics for the behavioral sciences*. Cengage Learning.

Greco, A., Annovazzi, C., Palena, N., Camussi, E., Rossi, G., & Steca, P. (2022). Self-efficacy beliefs of university students: Examining factor validity and measurement invariance of the new academic self-efficacy scale. *Frontiers in Psychology*, *12*, 498824 https://doi.org/10.3389/fpsyg.2021.498824

Grøtan, K., Sund, E. R., & Bjerkeset, O. (2019). Mental health, academic self-efficacy and study progress among college students – the SHoT study, Norway. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.00045

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. A. (2019). *Multivariate Data Analysis*

(8th ed.). Cengage Learning EMEA.
Hawe, E., Lightfoot, U., & Dixon, H. (2017). First-year students working with exemplars: promoting self-efficacy, self-monitoring and self-regulation. *Journal of Further and Higher Education*, 1–15. https://doi.org/10.1080/0309877X.2017.1349894
Hayat, A. A., Shateri, K., Amini, M., & Shokrpour, N. (2020). Relationships between academic self-efficacy, learning-related emotions, and metacognitive learning strategies with academic performance in medical students: a structural equation model. *BMC Medical Education*, *20*(1), 76. https://doi.org/10.1186/s12909-020-01995-9
Hinton, P. R., McMurray, I., & Brownlow, C. (2004). *SPSS explained*. Routledge.
Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118
Iacobucci, D. (2010). Structural equations modeling: Fit Indices, sample size, and advanced topics. *Journal of Consumer Psychology*, *20*(1), 90–98. https://doi.org/10.1016/j.jcps.2009.09.003
Ifdil, I., Bariyyah, K., Dewi, A. K., & Rangka, I. B. (2019). The college academic self-efficacy scale (CASES): An Indonesian validation to measure the self-efficacy of students. *Jurnal Kajian Bimbingan dan Konseling*, *4*(4), 115–121. https://doi.org/10.17977/um001v4i42019p115
Kline, R. B. (2023). *Principles and practice of structural equation modeling*. Guilford Publications.
Korgan, C., Durdella, N., & Stevens, M. (2013). The development of academic self-efficacy among first-year college students in a comprehensive public university. *Higher Education in Review*, *10*, 11–37. https://bpb-us-e1.wpmucdn.com/sites.psu.edu/dist/b/36443/files/2016/02/Vol10_Korganetal.pdf
Meyer, J. P. (2014). *Applied measurement with jMetrik*. Routledge.
Misra, R., & Castillo, L. G. (2004). Academic stress among college students: Comparison of American and International students. *International Journal of Stress Management*, *11*(2), 132–148. https://doi.org/10.1037/1072-5245.11.2.132
Morales-Rodríguez, F. M., & Pérez-Mármol, J. M. (2019). The role of anxiety, coping strategies, and emotional intelligence on general perceived self-efficacy in university students. *Frontiers in Psychology*, *10*, Article 443080. https://doi.org/10.3389/fpsyg.2019.01689
Mundfrom, D. J., Shaw, D. G., & Ke, T. L. (2005). Minimum sample size recommendations for conducting factor analyses. *International Journal of Testing*, *5*(2), 159–168. https://doi.org/10.1207/s15327574ijt0502_4
Muthén, B., & Muthén, L. (2017). Mplus. In W. J. van der Linden (Ed.), *Handbook of item response theory* (pp. 507–518). Chapman and Hall/CRC.
Nunally, J. (1978). *Psychometric methods* (2nd ed.). McGraw-Hill.
Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.
Owen, S. V., & Froman, R. D. (1988). *Development of a college academic self-efficacy scale*. National Council on Measurement in Education.
The Jamovi Project (2022). Jamovi (Version 2.3) [Computer Software]. https://www.jamovi.org
Rae, G. (2007). A note on using stratified alpha to estimate the composite reliability of a test composed of interrelated nonhomogeneous items. *Psychological Methods*, *12*(2), 177–184. https://doi.org/10.1037/1082-989X.12.2.177
Raykov, T., & Marcoulides, G. A. (2011). *Introduction to psychometric theory*. Routledge.
Rini, Q. K., Majorsy, U., & Hapsari, R. M. (2015). Hubungan metakognisi, efikasi diri akademik dan prestasi akademik pada mahasiswa. *Prosiding PESAT*, 66–71.

https://ejournal.gunadarma.ac.id/index.php/pesat/article/view/1340

Sagone, E., & Caroli, M. E. De. (2014). Locus of control and academic self-efficacy in university students: The effects of self-concepts. *Procedia - Social and Behavioral Sciences*, *114*, 222–228. https://doi.org/10.1016/j.sbspro.2013.12.689

Schunk, D. H., & Pajares, F. (2002). *The development of academic self-efficacy*. Allan Wigfield.

Taber, K. S. (2018). The use of Cronbach's Alpha when developing and reporting research instruments in science education. *Research in Science Education*, *48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Urbina, S. (2014). *Essentials of psychological testing*. Wiley. https://doi.org/10.1002/9781394259458

van Rooij, E. C. M., Jansen, E. P. W. A., & van de Grift, W. J. C. M. (2018). First-year university students' academic success: the importance of academic adjustment. *European Journal of Psychology of Education*, *33*(4), 749–767. https://doi.org/10.1007/s10212-017-0347-8

van Zyl, L. E., Klibert, J., Shankland, R., See-To, E. W. K., & Rothmann, S. (2022). The general academic self-efficacy scale: Psychometric properties, longitudinal invariance, and criterion validity. *Journal of Psychoeducational Assessment*, *40*(6), 777–789. https://doi.org/10.1177/07342829221097174

Widhiarso, W., & Ravand, H. (2014). Estimating reliability coefficient for multidimensional measures: A pedagogical illustration. *Review of Psychology*, *21*(2), 111–121. https://psycnet.apa.org/record/2016-00133-001

Xia, Y., & Yang, Y. (2019). RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. *Behavior Research Methods*, *51*(1), 409–428. https://doi.org/10.3758/s13428-018-1055-2

Zajacova, A., Lynch, S. M., & Espenshade, T. J. (2005). Self-efficacy, stress, and academic success in college. *Research in Higher Education*, *46*(6), 677–706. https://doi.org/10.1007/s11162-004-4139-z

Zysberg, L., & Schwabsky, N. (2021). School climate, academic self-efficacy and student achievement. *Educational Psychology*, *41*(4), 467–482. https://doi.org/10.1080/01443410.2020.1813690

Appendix A.
Introduction

Table A1
Review of the Previous Academic Self-Efficacy Measurements

Measurements	Participants	Psychometric property		Limitation
		Validity	Reliability	
The College Academic Self-Efficacy Scale				
English Version	Undergraduate students, in general	Concurrent Validity and Construct validity (Exploratory factor analysis)	Cronbach's alpha	- The EFA results do not display the estimated loading factor and goodness of fit index (Hair et al., 2019). - Cronbach's alpha is a reliability test for unidimensional assumptions, and stratified alpha is a reliability test for multidimensional constructs (Widhiarso & Ravand, 2014).
Indonesian Version	Students in general	None	Cronbach's alpha	- The requirements for participants—whether graduate, undergraduate, or secondary school students—are not specified. - Internal consistency for reliability test findings exists, but construct validity—whether through factor analysis or convergent and divergent validity—has not been thoroughly examined (Hair et al., 2019). - Widhiarso and Ravand (2014) state that reliability testing with Cronbach's alpha is more suited for unidimensional assumptions, whereas stratified alpha is better suited for multidimensional constructs.

Table A1. *Continued.*

The Academic Self-Efficacy Scale				
English Version	Italian university students in general	Construct validity (Exploratory factor analysis)	Cronbach's alpha	- All participants were Italian undergraduates studying law, psychology, and medicine. - The goodness of fit index and estimated factor loadings were not identified because the construct validity test employs EFA, which yields four components (Hair et al., 2019). - Cronbach's alpha, which is more suitable for unidimensional assumptions, is used in the reliability test; stratified alpha is more suitable for multidimensional constructs (Widhiarso & Ravand, 2014).
Indonesian Version	Indonesian university students, in general	Construct validity (Second-order confirmatory factor analysis)	Cronbach's alpha	- The construct validity test using CFA excluded three items. The CFA test findings show that all items are valid, although some goodness-of-fit indices are still poor (Hair et al., 2019). - Cronbach's alpha reliability tests are for unidimensional assumptions, whereas stratified alpha tests are for multidimensional constructs (Widhiarso & Ravand, 2014).

Table A1. *Continued.*

General Academic Self-Efficacy	University students in the USA and Netherlands	Construct validity (Unidimensional confirmatory factor analysis)	Cronbach's alpha	- The participants are undergraduate students without specification.
Academic Situation-Specific Perceived Self-Efficacy Scale	Latin American university students	Construct validity (Unidimensional confirmatory factor analysis)	Cronbach's alpha	- Latin American females made up the majority of the participants. - ASSPSES is not exclusively for first-year undergraduate students.
New Academic Self-Efficacy	Italian university students in general	Construct validity (Exploratory factor analysis and confirmatory factor analysis)	Cronbach's alpha and McDonald's Omega	- Participants are Italian undergraduate students who met certain requirements. - Stratified alpha and McDonald's hierarchical omega are relevant for multidimensional construct (Widhiarso & Ravand, 2014).

Appendix B.
Five Stages of Item Analysis

Table B1
Stage 1 - 52-Item TIASS on Second-order Confirmatory Factor Analysis

Items	Factor 1	Factor 2	Factor 3	Factor 4	Identification
Item1	0.466				Eliminated
Item2	0.547				Admitted
Item3	0.412				Eliminated
Item4 (1)	0.630				Admitted
Item5	0.468				Eliminated
Item6 (2)	0.587				Admitted
Item7	0.340				Eliminated
Item8 (3)	0.628				Admitted
Item9	0.313				Eliminated
Item10 (4)	0.562				Admitted
Item11 (5)	0.638				Admitted
Item12	0.282				Eliminated
Item13	0.381				Eliminated
Item14		0.420			Eliminated
Item15		0.301			Eliminated
Item16		0.366			Eliminated
Item17 (6)		0.535			Admitted
Item18 (7)		0.561			Admitted
Item19 (8)		0.619			Admitted
Item20		0.253			Eliminated
Item21 (9)		0.650			Admitted
Item22		0.404			Eliminated
Item23		0.386			Eliminated
Item24		0.432			Eliminated
Item25		0.494			Eliminated
Item26		0.469			Eliminated
Item27			0.343		Eliminated
Item28 (10)			0.608		Admitted
Item29 (11)			0.597		Admitted
Item30			0.215		Eliminated
Item31 (12)			0.587		Admitted
Item32 (13)			0.545		Admitted
Item33			0.421		Eliminated
Item34 (14)			0.642		Admitted
Item35			0.207		Eliminated
Item36 (15)			0.642		Admitted
Item37			0.253		Eliminated
Item38			0.502		Eliminated
Item39 (16)			0.621		Admitted
Item40				0.507	Eliminated
Item41				0.346	Eliminated
Item42				0.247	Eliminated
Item43				0.546	Eliminated
Item44				0.237	Eliminated
Item45 (17)				0.621	Admitted
Item46				0.583	Eliminated
Item47				0.084	Eliminated

Table B1. *Continued.*

Item48	-0.070	Eliminated
Item49	0.217	Eliminated
Item50 (18)	0.774	Admitted
Item51 (19)	0.772	Admitted
Item52 (20)	0.513	Admitted

Table B2

Stage 2 - 25-Item TIASS on Second-order Confirmatory Factor Analysis

Items	Factor 1	Factor 2	Factor 3	Factor 4	Identification
Item2	0.484				Eliminated
Item4 (1)	0.677				Admitted
Item6 (2)	0.551				Admitted
Item8 (3)	0.643				Admitted
Item10 (4)	0.584				Admitted
Item11 (5)	0.672				Admitted
Item17 (6)		0.531			Admitted
Item18 (7)		0.592			Admitted
Item19 (8)		0.656			Admitted
Item21 (9)		0.616			Admitted
Item28 (10)			0.583		Admitted
Item29 (11)			0.573		Admitted
Item31 (12)			0.610		Admitted
Item32 (13)			0.564		Admitted
Item34 (14)			0.650		Admitted
Item36 (15)			0.628		Admitted
Item38			0.484		Eliminated
Item39 (16)			0.605		Admitted
Item40				0.493	Eliminated
Item43				0.518	Eliminated
Item45 (17)				0.557	Admitted
Item46				0.617	Eliminated
Item50 (18)				0.806	Admitted
Item51 (19)				0.804	Admitted
Item52 (20)				0.511	Admitted

Table B3

Stage 3 - 22-Item TIASS on Second-order Confirmatory Factor Analysis

Items	Factor 1	Factor 2	Factor 3	Factor 4	Identification
Item4 (1)	0.681				Admitted
Item6 (2)	0.533				Admitted
Item8 (3)	0.620				Admitted
Item10 (4)	0.611				Admitted
Item11 (5)	0.696				Admitted
Item17 (6)		0.531			Admitted
Item18 (7)		0.595			Admitted
Item19 (8)		0.652			Admitted
Item21 (9)		0.616			Admitted
Item28 (10)			0.580		Admitted
Item29 (11)			0.588		Admitted
Item31 (12)			0.609		Admitted
Item32 (13)			0.539		Admitted
Item34 (14)			0.662		Admitted

Table B3. *Continued.*

Item36 (15)	0.637		Admitted
Item39 (16)	0.587		Admitted
Item43		0.485	Eliminated
Item45 (17)		0.527	Admitted
Item46		0.594	Eliminated
Item50 (18)		0.836	Admitted
Item51 (19)		0.833	Admitted
Item52 (20)		0.514	Admitted

Table B4

Stage 4 - 21-Item TIASS on Second-order Confirmatory Factor Analysis

Items	Factor 1	Factor 2	Factor 3	Factor 4	Identification
Item4 (1)	0.681				Admitted
Item6 (2)	0.533				Admitted
Item8 (3)	0.620				Admitted
Item10 (4)	0.611				Admitted
Item11 (5)	0.697				Admitted
Item17 (6)		0.533			Admitted
Item18 (7)		0.593			Admitted
Item19 (8)		0.649			Admitted
Item21 (9)		0.620			Admitted
Item28 (10)			0.580		Admitted
Item29 (11)			0.588		Admitted
Item31 (12)			0.606		Admitted
Item32 (13)			0.538		Admitted
Item34 (14)			0.663		Admitted
Item36 (15)			0.638		Admitted
Item39 (16)			0.590		Admitted
Item45 (17)				0.562	Admitted
Item46				0.481	Eliminated
Item50 (18)				0.852	Admitted
Item51 (19)				0.871	Admitted
Item52 (20)				0.519	Admitted

Table B5

Stage 5 - 20-Item TIASS on Second-order Confirmatory Factor Analysis

Items	Factor 1	Factor 2	Factor 3	Factor 4	Identification
Item4 (1)	0.682				Admitted
Item6 (2)	0.533				Admitted
IItem8 (3)	0.619				Admitted
Item10 (4)	0.611				Admitted
Item11 (5)	0.697				Admitted
Item17 (6)		0.534			Admitted
Item18 (7)		0.593			Admitted
Item19 (8)		0.647			Admitted
Item21 (9)		0.621			Admitted
Item28 (10)			0.580		Admitted
Item29 (11)			0.589		Admitted
Item31 (12)			0.605		Admitted
Item32 (13)			0.537		Admitted
Item34 (14)			0.663		Admitted
Item36 (15)			0.639		Admitted

Table B5. *Continued.*

Item39 (16)	0.590		Admitted
Item45 (17)		0.547	Admitted
Item50 (18)		0.841	Admitted
Item51 (19)		0.900	Admitted
Item52 (20)		0.525	Admitted