The General Academic Self-Efficacy Scale: Psychometric Properties, Longitudinal Invariance and Criterion Validity

The General Academic Self-Efficacy Scale: Psychometric Properties, Longitudinal Invariance and Criterion Validity

Llewellyn E. van Zyl[1,2,3,4*], Jeff Klibert[5], Rebecca Shankland[6], Eric W.K. See-To[7] & Sebastiaan Rothmann[2]

[1] University of Eindhoven, Human Performance Management, the Netherlands
[2] Optentia Research Unit, North-West University, South Africa
[3] Department of Human Resource Management, University of Twente, the Netherlands
[4] Department of Social Psychology, Goethe University, Frankfurt am Main, Germany
[5] Department of Psychology, Georgia Southern University, USA
[6] Laboratoire DIPHE, Université Lumière Lyon 2, France
[7] Department of Computing and Decision Sciences, Lingan University, Hong Kong

Abstract

Academic self-efficacy (ASE) refers to a student's global belief in his/her ability to master the various academic challenges at university and is an essential antecedent of wellbeing and performance. The five-item General Academic Self-Efficacy Scale (GASE) showed promise as a short and concise measure for overall ASE. However, of its validity and reliability outside of Scandinavia is limited. Therefore, this paper aimed to investigate the psychometric properties, longitudinal invariance, and criterion validity of the GASE within a sample of university students (Time 1: *n*=1056 & Time 2: *n*=592) in the USA and Western Europe. The results showed that a unidimensional factorial model of overall ASE fitted the data well, was reliable and invariant across time. Further, criterion validity was established by finding a positive relationship with task performance at different time stamps. Therefore, the GASE can be used as a valid and reliable measure for general ASE.

Keywords: Academic Self-Efficacy; Task Performance; Longitudinal Invariance; Longitudinal Confirmatory Factor Analysis; University Students

Abstract Word Limit: 150 (Current: 147)

Article Word Limit: 2200 (Current: 2322)

Van Zyl, L.E., Klibert, J., Shankland, R., See-To, E.W.K., & Rothmann, S. (Accepted). The General Academic Self Efficacy Scale: Psychometric Properties, Longitudinal Invariance and Criterion Validity. *Journal of Psychoeducational Assessment*.

INTRODUCTION

General academic self-efficacy (ASE) refers to students' global belief in their ability to master the various academic challenges at university and is an essential antecedent of wellbeing and academic performance (Nielson et al., 2018). Within university contexts, higher levels of ASE has been associated with lower levels of depression/stress/anxiety (Tahmassian & Jalali-Moghadam, 2011), better decision making, motivation, and engagement, as well as higher levels of academic- and task performance (Doo & Bonk, 2020; Tossavainen et al., 2021; Van Zyl et al., 2021). As a social-cognitive process, ASE is concerned with developing the belief in one's ability to obtain and optimize the cognitive, behavioural, emotional and social resources required to perform better at academic-related tasks (Nielson et al., 2018). A meta-analysis showed that ASE is the strongest predictor of overall performance at university i.t.o grade point average (over and above personality, motivation and learning styles) (Richardson et al., 2012). Various studies have also found that ASE is a strong predictor of students overall task performance (i.e. the proficiency to perform well in academic tasks through making the right choices and to take the initiative to perform the most important or core tasks central to their academic studies on time, and to specification) (Campbell & Hackett, 1986; Lim & Bang, 2018; Tossavainen et al., 2021). When students feel competent in their own academic abilities, they are better able to utilize their capabilities to prioritize the completion of competing academic tasks more effectively, are less likely to be discouraged by setbacks, less likely to procrastinate, and invest more effort into their studies (Richardson et al., 2012; Tossavainen et al., 2021). Given ASE's importance for performing well at academic tasks, it is not surprising that its development has become a central strategy for universities to enhance academic throughput (Meintjes, 2020).

As such, various psychometric instruments to measure ASE have been developed (cf. Dever & Kim, 2016; Lindstrøm & Sharma, 2011; Owen & Froman, 1998; Zimmerman et al., 1992). However, these instruments are exceptionally lengthy (ranging from 16 to 33) and have shown different factorial structures and varying ranges of internal consistency in different settings (Meintjes, 2020). This could lead to biased results and limits the potential for cross-cultural comparisons. Developed and validated in Scandinavia, the English adapted five-item General Academic Self-Efficacy Scale (GASE) showed promise as a short, clear and concise measure for overall academic self-efficacy (Nielson et al., 2018). The scale measures the global belief in one's ability to perform and plan tasks associated with an academic degree (Nielson et al., 2018). The GASE has shown to be a valid and reliable measure in various studies and proved to be invariant between genders (Bass, 2020; Hitches et al., 2021; Nielson, 2020). However, its validity and reliability outside of Scandinavia are yet to be investigated.

Therefore, the purpose of this paper was to investigate the psychometric properties, longitudinal invariance, and criterion validity (i.r.o. task performance) of the GASE within the tertiary educational environment in Western Europe and the US. The study aims to provide researchers and practitioners with evidence that the GASE can be used as a valid and reliable tool to measure general academic self-efficacy in university contexts.

METHODOLOGY

Research Design

A longitudinal, electronic survey-based research design was employed to explore the psychometric properties, longitudinal invariance and criterion validity of the GASE. Data for this paper forms part of a larger cross-cultural student wellbeing project obtained at two time-points over three months.

Participants

A purposive sampling strategy was employed to gather data from university students from one academic university in the USA and -Belgium as well as one technical university in the Netherlands. At Time 1 the majority of the 1056 participants were English speaking (48.3%), American (46.8%)[1] females (54.9%) between the ages of 21 and 25 (58.80%) who held part-time employment (55.9%). At Time 2, the majority of the 592 matched participants were Dutch-speaking (71.3%), Western European (66.4%) males (60.8%) between the ages of 21 and 25 (72.6%) who held part-time employment (62.8%).

[INSERT TABLE 1 HERE]

Measuring Instruments

The *General Academic Self-Efficacy scale* (GASE: Nielsen et al., 2018) was used to measure academic self-efficacy. This five-item self-report scale measures academic self-efficacy on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). An example item is: “I know I can pass the exam if I put in enough work during the semester”. Akanni and Oduaran (2019) reported acceptable levels of internal consistency with a Cronbach’s alpha of 0.81.

The *Task Performance sub-scale* of Koopmans et al.’s (2012) Individual Work Performance Scale was used to measure overall task performance. Task performance was measured by seven items on a six-point Likert-scale ranging from 1 (‘Never’) to 6 (‘Always’). An example of item is: “I knew how to set the right priorities”. In Europe, the scale also produced acceptable levels of internal consistency as represented by a McDonald’s Omega of 0.84 (Van Zyl et al., in press).

Statistical Analysis

[1] We grouped those who reported Dutch, Belgian, German, and French nationalities together as ‘Western European’. Those who reported other nationalities (e.g., Polish, Lithuanian, Greek) as “Other European”. Similarly, for participants from the USA who reported their nationalities to be Canadian, Mexican or Brazilian we grouped together as ‘Other American’.

Data were processed with JASP v. 0.15 (JASP, 2021) and Mplus v 8.6 (Muthén & Muthén, 2021) through Structural Equation Modelling (SEM) with the maximum likelihood estimator. The full maximum likelihood estimation method (FIML) was used to manage missing data.

First, a unidimensional confirmatory factor analytical (CFA) model was estimated for the scale at each time-point. Model fit was evaluated through conventional standards (c.f. Table 2).

[INSERT TABLE 2HERE]

Second, a longitudinal CFA (LCFA) approach was employed to determine the temporal stability of the scale's factor structure. Here, academic self-efficacy at Time 1 was regressed on academic self-efficacy at Time 2. The model had to show (a) good data-model fit (c.f Table 2), (b) excellent measurement quality (λ>0.40; p<0.01) (c) a positive regression path between the factors (p<0.01), and (d) the scale at Time 1 needed to explain at least 40% of the variance in Time 2 (Wong & Wong, 2020). Further, the average variance explained and reliability estimates (Cronbach Alpha > 0.70; McDonald's omega> 0.70) were computed (Wong & Wong, 2020).

Third, longitudinal measurement invariance (LMI) was used to determine the configural (similar factor structure), metric (similar factor loadings), and scalar invariance (similar intercepts) of the scale over time. Invariance was established by comparing these models based on the following criteria: changes in RMSEA (Δ < 0.015), SRMR (Δ < 0.02 for configural vs. metric/scalar; Δ < 0.01 metric vs. scalar), CFI (Δ < 0.01) and TLI (Δ < 0.01) (Wong & Wong, 2020). Differences in χ^2 were not considered (but reported for transparency) due to its sensitivity to both sample size and model complexity (Morin et al., 2020; Van Zyl & ten Klooster, 2022).

Finally, criterion validity was established by estimating concurrent and predictive validity via a structural model. Concurrent validity was estimated by relating Academic Self-Efficacy at Time 1 to Task Performance at Time 1 and Academic Self-Efficacy at Time 2 to Task Performance at Time 2. To establish predictive validity, Academic Self-Efficacy at Time 1 was regressed on Task Performance at Time 2. To enhance fit, Item 1 and Item 2 of the Task Performance scale were permitted to correlate. For each regression, a significance level of p < 0.01 was set.

RESULTS

Confirmatory Factor Analysis: Cross-Sectional and Longitudinal

The factorial validity of the GASE at each time point were explored through estimating a single first-order factor model. Here all items were specified to load directly onto a single factor. Observed items were used as indicators for the latent factor. No items were removed. The results summarised in Table

3 showed that the model fitted the data well at both Time 1 ($\chi^2_{(1056)}$ = 17.73 p>0.001; df = 5; CFI = 0.99; TLI = 0.98; RMSEA = 0.05 [.026, .075]; SRMR = 0.02; AIC= 11905.73; BIC = 11980.16) and Time 2 ($\chi^2_{(592)}$ = 12.24 p>0.001; df = 5; CFI = 0.99; TLI = 0.98; RMSEA = 0.05 [.014, .085]; SRMR = 0.02; AIC= 6673.08; BIC = 6738.83).

[INSERT TABLE 3 HERE]

Thereafter, a LCFA model was estimated. Here ASE at Time 1 was regressed on ASE at Time 2. Error variances between time-points were permitted to covary. Table 3 indicates that this model also fitted the data well ($\chi^2_{(1056)}$ = 45.60 p>0.001; df = 29; CFI = 0.99; TLI = 0.99; RMSEA = 0.02 [.008, .036]; SRMR = 0.02; AIC= 18015.40; BIC = 18194.04). GASE at Time 1 was also significantly related to GASE at Time 2 (β = 0.69; S.E = 0.03; t-value= 22.05; p < 0.01) and explained 47% of the overall variance.

Item Level Parameter Estimates and Internal Consistency

Table 4 summarizes the item level parameter estimates and internal consistency of the LCFA model. All factor loadings were significant and greater than 0.40 at both time-points. The AVE was 40% at Time 1 and 45% at Time 2. The scale showed acceptable levels of internal consistency at both Time Points with Cronbach's alpha and McDonald's Omega ranging from 0.74 to 0.78.

[INSERT TABLE 4 HERE]

Longitudinal Measurement Invariance and Mean Comparisons

Measurement equivalence was investigated through LMI. The results summarised in Table 5 showed that metric, configural and scalar invariance was established. No significant differences in terms of RMSEA (Δ < 0.015), SRMR (Δ < 0.02 for configural vs. metric/scalar; Δ < 0.01 metric vs. scalar), CFI (Δ < 0.01) and TLI (Δ < 0.01) between the configural, metric, and scalar invariance models were found (Wong & Wong, 2020). Therefore, the GASE showed to be a consistent measure over time and mean comparisons can be made. Latent Mean comparisons, with GASE at Time 1 set as the reference point, showed that GASE decreased slightly over three months ($\Delta \bar{x}$ = -0.20; S.E = 0.04; p =0.00).

[INSERT TABLE 5 HERE]

Concurrent and Predictive Validity

Finally, separate structural models were estimated to evaluate the concurrent and predictive validity of the instrument (c.f. Table 6). Concurrent validity was established through finding a positive relationship between GASE and Task Performance at both Time 1 ($\chi^2_{(1056)}$ = 215.56; df = 52; CFI = 0.97; TLI =

0.96; RMSEA = 0.06 [.047, .062]; SRMR = 0.03; AIC= 32430.12; BIC = 32618.68; β= 0.54; S.E= 0.03; R^2= 0.29) and Time 2 ($\chi^2_{(592)}$ = 210.06; *df* = 52; CFI = 0.94; TLI = 0.92; RMSEA = 0.07 [.062, .082]; SRMR = 0.04; AIC= 18412.02; BIC = 18578.59; β= 0.58; S.E= 0.04; R^2= 0.34).

[INSERT TABLE 6 HERE]

Predictive validity was also established through finding a positive relationship between GASE at Time 1 with Task Performance at Time 2 ($\chi^2_{(1056)}$ = 174.633; *df* = 52; CFI = 0.96; TLI = 0.95; RMSEA = 0.05 [.040, .055]; SRMR = 0.04; AIC= 23741.36; BIC = 23929.93; *β*= 0.39; S.E= 0.08; R^2= 0.15).

DISCUSSION

The purpose of this paper was to investigate the psychometric properties, longitudinal invariance, and criterion validity (i.t.o. task performance) of the GASE within a US and Western European tertiary educational environment. The results showed that a single, first-order factorial model of overall academic self-efficacy fitted the data well, was reliable and invariant across time. In line with Nielson et al. (2018), our findings show that the GASE measures general academic self-efficacy validly and reliably. Therefore, the mean scores of the GASE could be used by educational practitioners to measure ASE and track the effectiveness of educational programmes or interventions aimed at enhancing ASE over time.

Further, criterion validity was established by finding a positive relationship with task performance at different time stamps. The results imply that when individuals hold active and positive beliefs about their abilities to plan/perform certain educational tasks or manage the challenges associated with their study programmes, they are more likely to perform better in their academic-related tasks. This is because the feeling that one has mastery over the skills required to perform a given educational task enhances the engagement and motivation required to perform (Tossavainen et al., 2021). Further, according to Richardson et al. (2012) this could also be because holding a high level of ASE affects how obstacles or challenges are viewed (opportunities to learn vs setbacks) which in turn leads to sustained task-related performance over time (Richardson et al., 2012; Tossavainen et al., 2021).

In conclusion, our results support the GASE as a valid and reliable measure for general academic self-efficacy within the current context. However, the study has its limitations. First, the sample is limited to a single US and two Western European universities. The results may therefore not be generalizable. Second, only (self-reported) task performance was used to establish criterion validity. Future research should include objective performance measures (e.g., grades) and other metrics associated with ASE, such as engagement, motivation, and resilience. Second, although various mechanisms were implemented to manage potential sample size attrition over time (e.g. students obtained course credit for participation; the follow-up assessment was kept as brief as possible; multiple reminders being sent: Mason, 1999) there was a 44% dropout between Time 1 and Time 2. This could have led to attrition bias which may affect the internal validity of the LFA and LMI assessments. However, the sample size at Time 2 is large enough to capture a full range of variation in responses, and therefore the configural, metric and scalar invariance assessments are valid for the current study. Future research should attempt to manage the dropout rate through implementing more implicit and explicit incentives for participation at Time 2. Finally, with increased global competition between academic institutions, potential students and future employers may be interested in how effective academic programmes are in developing more self-efficacious students. Given that the invariance between the two nations was not investigated due to sample size limitations, these types of comparisons can not be made. Future research should aim to test the measurement equivalence of the scale across cultures.

REFERENCES

Akanni, A. A., & Oduaran, C. A. (2018). Perceived social support and life satisfaction among freshmen: Mediating roles of academic self-efficacy and academic adjustment. *Journal of Psychology in Africa*, *28*(2), 89-93.

Bass, K. L. (2020). *Can Intentional Instruction Influence Self-efficacy and Increase Academic Performance? A Pygmalion Effect in Spanish Language Learners* (Doctoral dissertation, Texas A&M University-Central Texas).

Campbell, N. K., & Hackett, G. (1986). The effects of mathematics task performance on math self-efficacy and task interest. *Journal of Vocational Behavior*, *28*(2), 149-162.

Curran, P. J., Bollen, K. A., Chen, F., Paxton, P., & Kirby, J. B. (2003). Finite sampling properties of the point estimates and confidence intervals of the RMSEA. *Sociological Methods & Research*, 32, 208-252.

Dever, B. V., & Kim, S. Y. (2016). Measurement equivalence of the PALS academic self-efficacy scale. *European Journal of Psychological Assessment*, *32*(1), 61-67

Doo, M. Y., & Bonk, C. J. (2020). The effects of self-efficacy, self-regulation and social presence on learning engagement in a large university class using flipped Learning. *Journal of Computer Assisted Learning*, *36*(6), 997-1010.

Hitches, E., Woodcock, S., & Ehrich, J. (2021). Shedding light on students with support needs: Comparisons of stress, self-efficacy, and disclosure. *Journal of Diversity in Higher Education*.

Hu, L. T., & 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.

JASP. (2021). Jeffery's Amazing Stats Program (v.0.15). [Statistical Software]. Accessed via the World Wide Web: https://jasp-stats.org

Kaiser, H. F., & Rice, J. (1974). Little jiffy, mark IV. *Educational and psychological measurement, 34*(1), 111-117.

Koopmans, L., Bernaards, C. M., Hildebrandt, V. H., Van Buuren, S., Van der Beek, A. J., & De Vet, H. C. W. (2012). Development of an individual work performance questionnaire. International *Journal of Productivity and Performance Management*, *62*(1), 6–28. doi:10.1108/17410401311285273

Lim, H. A., & Bang, E. J. (2018). The effects of music listening on affect, self-efficacy, mental exertion, and task performance of online learners. *Journal of the Scholarship of Teaching and Learning for Christians in Higher Education*, *8*(1), 5-19.

Lindstrøm, C., & Sharma, M. D. (2011). Self-efficacy of first year university physics students: do gender and prior formal instruction in physics matter?. International Journal of Innovation in Science and Mathematics Education, 19(2), 1-19.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130-149.

Mason, M. (1999). A review of procedural and Statistical methods for handling attrition and missing data in clinical research. Measurement and Evaluation in Counseling and Development, 32, 111–118.

Meintjes, R. (2020). *Depression and self-regulated learning as predictors of first-year students' academic performance: a case study* (Doctoral dissertation, University of the Free State).

Morin, A.J.S., Myers, N.D., & Lee, S. (2020). Modern factor analytic techniques: bifactor models, exploratory Exploratory Structural Equation Modeling 19 structural equation modeling and bifactor-ESEM. In G. Tenenbaum, & Eklund, R.C. (Eds.), Handbook of Sport Psychology, 4th Edition, Vol. 2 (pp. 1044-1073). New York, NY: Wiley Publishers.

Muthén, L. K., & Muthén, B. O. (2021). *Mplus* (Version 8.4). [Statistical software]. Los Angeles, CA: Muthén and Muthén.

Nielsen, T. (2020). Measuring differences in Current Academic Learning Self-Efficacy (CAL-SE) and Current Academic Exam Self-Efficacy (CAE-SE) for social science students: a Rasch-based multi degree-program validity study. In *Research on Teaching and Learning in Higher Education* (p. 57). Waxmann Verlag.

Nielsen, T., Dammeyer, J., Vang, M. L., & Makransky, G. (2018). Gender fairness in self-efficacy? A Rasch-based validity study of the General Academic Self-efficacy scale (GASE). *Scandinavian Journal of Educational Research*, *62*(5), 664-681.

Owen, S. V., & Froman, R. D. (1988, April). Development of a College Academic Self-Efficacy Scale. Paper presented at the Annual Meeting of the National Council on Measurement in Education, New Orleans, LA.

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. Psychological bulletin, 138(2), 353–387. doi:10.1037/a0026838

Tahmassian, K., & Jalali-Moghadam, N. (2011). Relationship between self-efficacy and symptoms of anxiety, Depression, worry and social avoidance in a normal sample of students. Iranian Journal of Psychiatry and Behavioral Sciences, 5(2), 91–98.

Tossavainen, T., Rensaa, R. J., & Johansson, M. (2021). Swedish first-year engineering students' views of mathematics, self-efficacy and motivation and their effect on task performance. *International Journal of Mathematical Education in Science and Technology*, *52*(1), 23-33.

Tossavainen, T., Rensaa, R. J., & Johansson, M. (2021). Swedish first-year engineering students' views of mathematics, self-efficacy and motivation and their effect on task performance. *International Journal of Mathematical Education in Science and Technology*, *52*(1), 23-38.

Van Zyl, L.E., & Ten Klooster, P.M. (2022). Exploratory Structural Equation Modelling: Practical Guidelines and Tutorial with a Convenient Online Tool for Mplus. *Frontiers in Psychiatry, 12*:795672 doi:0.3389/fpsyt.2021.795672

Van Zyl, L.E., Van Vuuren, M., Roll, L.C., & Stander, M.W. (In review). Person-environment fit and task performance: The mediating role of Grit. *Current Psychology*.

Van Zyl, L.E., Rothmann, S., & Zondervan-Zwijnenburg, M.A.J. (2021). Longitudinal Trajectories of Study Characteristics and Mental Health before and during the COVID-19 Lockdown. Frontiers in Psychology, 12, 1-13. doi: 10.3389/fpsyg.2021.63353

Wong, J., & Wong, X. (2020). *Structural equation modelling: Applications using Mplus* (2nd ed.). Chichester, UK: Wiley & Sons.

Wu, H., Li, S., Zheng, J., & Guo, J. (2020). Medical students' motivation and academic performance: the mediating roles of self-efficacy and learning engagement. *Medical education online*, *25*(1), 1742964.

Table 1. *Demographic characteristics of participants at Time 1 (N=1056) and Time 2 (N=592)*

Item	Category	Time 1		Time 2	
		Frequency (*f*)	Percentage (%)	Frequency (*f*)	Percentage (%)
Gender	Male	473	44.80	360	60.8
	Female	580	54.90	230	38.9
	Other	3	0.30	2	0.3
Age (years)	18-20 years	370	35.00	128	21.6
	21-25 years	621	58.80	430	72.6
	26-30 years	36	3.40	18	3.0
	31 years and older	29	2.70	16	2.7
Nationality	European	408	38.60	393	66.4
	USA	494	46.80	65	11.0
	Other European	130	12.30	130	22.0
	Other American	24	2.30	4	0.70
Home Language	English	509	48.20	78	13.2
	Dutch	435	41.20	422	71.3
	Other	112	10.60	92	15.5
Employment	Full Time	50	4.70	13	2.2
	Part Time	590	55.9	372	62.8
	Unemployed	416	39.4	207	35.0

Table 2. *Model Fit Statistics*

Fit indices	Cut-Off Criterion	Sensitive to N	Penalty for Model Complexity
Absolute fit indices			
Chi-Square (χ2)	• Lowest comparative value between measurement models • Non-Significant Chi-Square value (p > 0.01)	Yes	No
Approximate Fit Indices			
Root-Means-Square Error of Approximation (RMSEA)[2]	• 0.06 to 0.08 (Marginally Acceptable); 0.00 to 0.05 (Excellent) • Non-significant RMSEA estimate (p > 0.05)	No	Yes
Standardized Root Mean Square Residual (SRMR)	• 0.06 to 0.08 (Marginally Acceptable); 0.01 to 0.05 (Excellent)	Yes	No
Incremental fit indices			
Comparative Fit Index (CFI)	• 0.90 to 0.95 (Marginally Acceptable Fit); 0.96 to 0.99 (Excellent)	No	No
Tucker-Lewis Index (TLI)	• 0.90 to 0.95 (Marginally Acceptable Fit); 0.96 to 0.99 (Excellent)	No	Yes
Akaike Information Criterion (AIC)	• Lowest value in comparative measurement models	Yes	Yes
Bayes Information Criterion (BIC)	• Lowest value in comparative measurement models	Yes	Yes

Adapted from Hu & Bentler (1999), Van Zyl & Ten Klooster (2022) and Wong & Wong (2020)

[2] RMSEA assesses the degree to which a hypothesized model differs from a 'perfect [hypothesised] model'. It acts a s supplementary fit measure to compensate for the sample size penalty imposed when using Chi-Square tests for model comparison. Three criteria are important for RMSEA estimates: a) the estimate range should be between 0.00 and 0.08, b) the RMSEA should differ significantly from the baseline model (i.e. the estimate should have a non-significant *p-value* (>0.01 or 0.05)) and c) in simpler models or extremely complex models (like bifactor- or ESEM models) the 90% CI range should not produce a negative value at the lower CI range indicator (Curran et al., 2003; Hu & Bentler, 1999; MacCallum et al., 1996; Morin et al., 2020; Van Zyl & Ten Klooster, 2022).

Table 3. *Cross-sectional and Longitudinal Confirmatory Factor Analyses*

Model	Type	χ^2	*df*	CFI	TLI	RMSEA		SRMR	AIC	BIC	aBIC	Meets Criteria
Cross-Sectional CFA												
Model 1	Unidimentional Model at Time 1	**17.73**	5	0.99	0.98	**0.05**	[.026-.075]	0.02	11905.73	11980.16	11932.52	Yes
Model 2	Unidimentional Model at Time 2	**12.24**	5	0.99	0.98	**0.05**	[.014-.085]	0.02	6673.08	6738.83	6691.21	Yes
Longitudinal CFA												
Model 3	Longitudinal Factor Analysis: Unidimentional Model	**45.60**	29	0.99	0.99	**0.02**	[.008-.036]	0.02	18015.40	18194.04	18079.70	Yes

χ^2 = Chi-square; *df* = degrees of freedom; TLI = Tucker-Lewis Index; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation [90%CI]; SRMR = Standardized Root Mean Square Residual; AIC = Akaike Information Criterion; BIC = Bayes Information Criterion; aBIC = Adjusted Bayes Information Criterion. **Bold:** Non-significant p >0.001

Table 4. *Item Level Parameter Estimates And Internal Consistency of the GASE*

Factor	Item Description	Time 1					Time 2				
		λ	S.E.	AVE	ω	α	λ	S.E.	AVE	ω	α
Academic Self-Efficacy				0.40	0.74	0.74			0.45	0.78	0.78
GASE_1	I generally manage to solve difficult academic problems if I try hard enough	0.69	0.02				0.70	0.03			
GASE_2	I know I can stick to my aims and accomplish my goals in my field of study	0.70	0.02				0.69	0.03			
GASE_3	I will remain calm in my exam because I know I will have the knowledge to solve the problems	0.59	0.03				0.66	0.03			
GASE_4	I know I can pass the exam if I put in enough work during the semester	0.67	0.02				0.72	0.03			
GASE_5	The motto ‘If other people can, I can too’ applies to me when it comes to my field of study.	0.47	0.03				0.55	0.03			

λ: Standardized factor loadings. S.E.: Standard Error. AVE: Average Variance Extracted. ω: McDonald’s Omega. α: Cronbach’s Alpha.

Table 5. *Longitudinal Invariance of the Unidimensional General Academic Self-Efficacy Scale*

No	Model	χ^2	*df*	CFI	TLI	RMSEA		SRMR	Model comparison	$\Delta\chi^2$	ΔCFI	ΔTLI	ΔRMSEA	ΔSRMR	Meets Invariance Criteria
M1	Configural Invariance	**45.60**	29	0.99	0.99	**0.023**	[.008-.036]	0.023	M3 vs M1	43.55	-0.01	-0.01	0.015	0.016	Yes
M2	Metric Invariance	**51.72**	32	0.99	0.99	**0.024**	[.011-.036]	0.034	M2 vs M1	6.12	0.00	0.00	0.001	0.011	Yes
M3	Scalar Invariance	**89.15**	35	0.98	0.98	**0.038**	[.029-.048]	0.039	M3 vs M2	37.43	-0.01	-0.01	0.014	0.005	Yes

χ^2= Chi-square; df = degrees of freedom; TLI = Tucker-Lewis Index; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation [90%CI]; SRMR = Standardised Root Mean Square Residual

Table 6. *Concurrent and Predictive Validity: General Academic Self-Efficacy and Task Performance*

Type	Regression Path			Standardized					
				Beta	S.E	t-value	*p*	R^2	Validity Established
Concurrent	Academic Self-Efficacy Time 1	→	Task Performance Time 1	0.54	0.03	18.82	0.00	0.29	Yes
	Academic Self-Efficacy Time 2	→	Task Performance Time 2	0.58	0.04	16.09	0.00	0.34	Yes
Predictive	Academic Self-Efficacy Time 1	→	Task Performance Time 2	0.39	0.05	8.62	0.00	0.15	Yes

→ = Regression; *β* = Standardized Beta; S.E = Standard Error; *p* = statistical significance; R^2 = Variance