Daengku: Journal of Humanities and Social Sciences Innovation *ISSN: 2775-6165 (online)*
Vol. 6 No. 4 (2026)
https://doi.org/10.35877/454RI.daengku4820

Exploring the Psychometrical Properties of Achievement Goals in Indonesian College Students

Siti Dini Fakhriya[a,*], Jesyia Meyrinda[b], Tsurayya Syarif Zain[c], & Melisa Kurnia Asfitri[c]

[a]*Universitas Islam Negeri Raden Fatah Palembang, Indonesia*
[b]*Institut Agama Islam Nadhalatul Ulama, Sumatera Selatan, Indonesia*
[c]*Politeknik Negeri Sriwijaya Palembang, Indonesia*

Abstract

Achievement Goals Theory (AGT) is focused on individuals' motivation and goal-pursuit behavior to achieve attainable targets. Therefore, this study aimed to investigate the validity and reliability of the AGT scale among Indonesian college students. Participants in this study included 220 students sampled using non-probability sampling. The Elliot's Achievement Goals Questionnaire-Revised (AGQ-R) was used to measure participants' beliefs about mastery-goal, mastery-avoidance, performance-goal, and performance-avoidance goals with three items on each subscale. Concurrent validity, Confirmatory Factor Analysis (CFA) and Rasch measurement model were carried out. Subsequently, Cronbach's Alpha, composite reliability were used to test reliability. The result showed that the three-factor model was the most appropriate representation as goodness of fit indexes $\chi 2$ (166) = 202.498, p = .0000, RMSEA = 0.045, 90% CI = 0.051− 4.493, probability RMSEA = 2.227, CFI =1.000, TLI = 1.000, SRMR = 0.056. AGQ-R was shown to be a valid and reliable instrument to be used on students in Indonesia. The three-factor model was the most appropriate representation, as well as a valid and reliable instrument to measure achievement goals in the sample of students in Indonesia. This model consisted of mastery and performance goals, while mastery-avoidance and performance-avoidance were combined into one domain. This validated psychometric instrument use to assess student motivation and design culturally-aligned interventions in the higher education and allow a development for teaching strategies that support student engagement and academic success.

Keywords: : Achievement Goal*;* AGQ-R*;* Confirmatory Factor Analysis*;* Indonesian*;* Psychometric*;* College Students

Received: 10 May 2026 Revised: 12 July 2026 Published: 31 August 2026

1. Introduction

Achievement Goals Theory (AGT) is focused on individuals' motivation and goal pursuit behavior toward attainable targets. The effectiveness of AGT depends on the organisms, particularly the thought and judgment, as well as the type of outcome scale utilized. It focuses on how individuals' cognitions about competence and goals shape the motivation (Urhahne & Wijnia, 2023). Previous studies have found evidence of the multidimensionality of achievement goals (Roberts and Nerstad, 2020; Cheng, 2023; Chazan et al., 2022). Students from Colombia, Spain, and Canada often interpret the achievements through a focus on mastering skills and a desire to avoid work. Consequently, there is a need to further investigate these assumptions, particularly in the context of college students in Africa. Previous studies in Namibia and Zimbabwe have shown that the existing models for understanding achievement goals may not be suitable for these populations (Becerra-Posada et al.2022; Tze et al., 2022; Kwon et al.2021).

In education, AGT frameworks have helped explain motivational factors influencing student persistence and achievement, considering both quantitative outcomes and socio-cultural influences (Jannini et al., 2024). AGT-derived scales have also been used in language education to validate motivation constructs in specific academic domains, improving instruction and learner support (Deeba & Ahmad, 2023). In physical education, motivational profiles based on AGT have informed pedagogical goals by identifying patterns associated with self-determination and effort (Boich et al., 2008).AGT distinguishes between mastery and performance goals, which focus on effort, skills, and normative

* Corresponding author.
E-mail address: sitidinifakhriya_uin@radenfatah.ac.id

performance comparisons (Jannini, Zeynep, & Meneke, 2024). Students who desire mastery and are focused on learning are probably intrinsically motivated. Meanwhile, students who are attempting merely to do better than others and to receive normatively based judgments of worth may be eristically motivated (Elliot, 1999). A previous study proposed the 2×2 framework that fully crossed competence and valence in the conceptualization of achievement goals (Elliot & McGregor, 2001). This four-factor model discusses students' achievement goals in mastery-goal, mastery-avoidance, performance-goal, and performance-avoidance goals (Apostolou, 2013; Elliot & Murayama, 2008).

Mastery goals primarily concern the development of personal competence through effort and learning, focusing on self-improvement and task mastery. Performance goals focus on normative standards of competence and the desire to show superiority or avoid negative judgments (Was et al., 2017). Performance-goal goals showcase an individual's abilities compared to peers. Mastery-avoidance goals reflect students' tendency to avoid situations that can lead to failure to learn or master tasks. Meanwhile, performance-avoidance goals included avoiding situations where competence might be negatively evaluated (Yeh et al., 2019). Several studies have used the Achievement Goals Questionnaire-Revised (AGQ-R) to examine the relationship between learning engagement, academic outcomes, and metacognitive strategies (Senko et al., 2011).

Elliot & Murayama (2008) developed AGQ-R, and it has been a measurement model for achievement goals. According to a study by Ning (2017), AGQ-R has sufficient psychometric qualities for use with Singaporean school-age children. Rosas (2015) found that there was a high level of internal consistency among the four achievement goals factors of AGQ-R for use with university students in Argentina. Another study confirmed a four-factor structure in AGQ-R. Internal consistency and test-retest reliability estimates were found to be satisfactory, construct- and discriminant-related evidence of validity were supported (Muis & Winne, 2012). The four-factor structure of the Thai version was confirmed for AGQ-R, according to the results of exploratory and Confirmatory Factor Analysis (CFA) (Ratsameemonthon, 2015). In a study of White and Black students, Campbell et al. (2008) found that AGQ-R provided the greatest fit to the data under the four-factor model and was consistent across both populations.

However, there are mixed results regarding the four-factor structure for this measurement model. Strunk (2014) on achievement goals presents inconsistent results after subjecting AGQ-R to CFA. The result showed that a three-factor solution provided the optimal fit in the sample of undergraduate students. Simamora & Mutiarawati (2021) also reported that the trichotomous model was the most effective model for new undergraduate students, according to CFA, and the 2x2 model showed the best fit for old students. Donnellan (2008) used exploratory methods to identify a three-factor framework that included learning objectives, performance goals, and outcome goals. According to Cook et al. (2017), AGQ-R scores suited a three-factor (rather than four-factor) model that reduced avoidance into one domain.

The different and mixed results raise concerns about the applicability of Western-developed models, such as AGQ-R, in non-Western and underrepresented populations. Cross-cultural studies have shown considerable variation in how achievement goals manifest, particularly between individualistic and collectivist cultures, and reported the influence of gender, learning environment, and classroom climate on goal orientations (Tze et al., 2022; Fokkema et al., 2024; Becerra-Posada et al., 2022; Kwon et al., 2021). Recent developments have also introduced broader constructs, such as personal-best and potential-based goals, and methodological innovations aimed at improving measurement validity and cultural sensitivity (Roberts & Nerstad, 2020; Chazan et al., 2022).

The validation study of AGQ-R in Indonesia has been conducted by Rodliya (2018). The results showed that AGQR was a valid tool for high school students, as evidenced by its high reliability with Cronbach's Alpha, Spearman-Brown reliability, and construct validity test (CFA). Another study conducted by Herdian (2021) also stated that the achievement goals orientation measuring instrument had good reliability and validity in students, but the coefficient value of the item correlation test was used to assess validity. Therefore, further study is needed regarding AGQ-R for undergraduate students in Indonesia.

The efficacy and generalizability of the study on achievement goals are critically dependent on the psychometric soundness of the instruments used. The observed inconsistencies in AGQ-R's factor structure across diverse samples, as evidenced by the varying results from (Chowning & Campbell, 2009; Rosas, 2015; Tomczak et al., 2024; Yeh et al., 2019), show a significant measurement challenge. These discrepancies suggest the urgent need for comprehensive psychometric evaluations, particularly concerning factor structure and measurement invariance, in new and underexplored populations. Without robust psychometric properties, the utility of the instrument for both advancing theoretical understanding and developing effective educational interventions is severely limited, showing the continuous need for validation in varied contexts. Therefore, this study aimed to expand the measurement framework

for AGQ-R by conducting factor analysis in a student population. This study aimed to explore the psychometric properties, including confirmatory factor and invariance analysis, of Achievement Goals in College Students.

2. Research Method and Materials

2.1. Participant

Participants were 220 undergraduate students from Palembang Indonesia who voluntarily participated after providing informed consent. According to Comrey and Lee (Williams et al., 2010), conducting CFA required a minimum of 200 participants, which was met in this study. Participants included 80 men (36.4%) and 140 women (63.6%), with a mean age of 19-20 years (M=3.48; SD=1.78). Data were collected using non-probability sampling with a convenience sampling method from undergraduate students in Palembang. All participants provided consent and completed the questionnaire online in June 2024. This study adopted a quantitative method, and participants voluntarily provided informed consent, confirming the willingness to participate

2.2. Research Instruments

Elliot's AGQ-R was used to measure participants' beliefs about mastery-goal, mastery-avoidance, performance-goal, and performance-avoidance goals with three items on each subscale. These instruments have been used widely in non-health professions, medical, and pharmacy education.

2.3. Descriptive Analysis

Descriptive analysis included the calculation of mean values, standard deviation, and normality testing. The average value (mean), standard deviation, and normalcy test were determined using the descriptive analysis through the IBM SPSS 2.9. Skewness and Kurtosis coefficients of ± 2 and ± 7 were used in normality testing (Hair, 2019). Central tendency and variability were shown by the mean and standard deviation (Gravetter et al., 2021).

2.4. Reliability Testing

The psychometric property tests used reliability and validity tests. The reliability test included Cronbach's Alpha, composite reliability, and the Rash measurement model. The validity test used concurrent validity and construct validity (Gravetter et al., 2021; Urbina, 2014). In this study, reliability was primarily assessed using Cronbach's Alpha, with values of $\alpha \geq .50$ considered adequate and .70–.90, showing high reliability (Kaplan, 2018). However, recognizing the limitation of Cronbach's Alpha to unidimensional measures (Cronbach, 1951), and considering the multidimensional nature of the instrument, stratified alpha was applied to obtain a more accurate reliability estimate. Stratified alpha, introduced by Cronbach et al. (1965), was specifically suited for multidimensional constructs as it accounted for each dimension's unique contribution. Following Rae's (2007) recommendations, Cronbach's Alpha was first calculated for each dimension, and the coefficients were integrated into the stratified alpha formula to provide a comprehensive reliability score for the entire factorial structure (Holzinger & Swineford, 1937; Kaplan, 2018).

2.5. Validity Testing

Validity referred to the extent to which evidence supported the interpretations and inferences made from test scores. Both construct and concurrent validity (Gravetter et al., 2021; Kaplan, 2018) were used in this study. Construct validity was primarily tested using CFA through Mplus version 7.0 (Muthén & Muthén, 2017). This method assessed how well the observed data fit the specified theoretical model, validating the factorial structure and internal consistency. Goodness-of-fit indicators (e.g., RMSEA $\leq .08$, SRMR $\leq .08$, CFI $\geq .95$, TLI $\geq .90$) were reviewed to verify the model's adequacy (Hair et al., 2019). Although chi-square ($\chi 2$) was reported, the sensitivity to sample size rendered it less critical for model fit evaluation (Clark & Bowles, 2018). The convergent validity was assessed by examining the Average Variance Extracted (AVE) value $\geq .50$ (Hair et al., 2019; Kaplan, 2018) showed that a construct explains more than half of the variance in its items. Similarly, an AVE below .50 suggested higher measurement error. This study used concurrent validity by correlating AGQ-R with other measures based on the same construct (Hair, 2019). Concurrent validity was carried out by testing the correlation coefficient of AGQ-R with other achievement goals measurements. In this case, the study used achievement goals measures, namely mastery, performance, mastery-

avoidance, and performance-avoidance goals, developed by Elliot & Murayama (2008).

3. Results and Discussion

The psychometric properties of the instruments utilized in this study were rigorously assessed through both reliability and validity tests.

Table 1. Descriptive Statistics, Convergent Validity of CFA Model

Dimensions	CR	AVE	α	Mean ± SD	Skewness	Kurtosis
Mastery Goal	0.746	0.721	0.617	2.809 ± 0.520	-0.087	-0.015
Performance Goal	0.739	0.675	0.629	2.955 ± 0.545	0.145	-0.280
Mastery Avoidance and Performance-Avoidance	0.869	0.778	0.749	3.025 ± 0.538	0.090	-0.312

Note: CR: composite reliability, AVE: Average Variance Extracted, α: Cronbach's Alpha coefficient

Table 1 shows the results of the reliability test and descriptive analysis. Referring to Hair et al. (2019), data distribution was declared normal when it ranged between the skewness coefficient ± 2 and the kurtosis coefficient ± 2. Therefore, each dimension of AGQ-R had normally distributed data. In more detail, interaction at mastery goal (M = 2.809, SD = 0.520) showed normally distributed data (Skewness = −0.087, Kurtosis = −0.015). Performance goal (M = 2.955, SD = 0.545) also showed a normal distribution (Skewness = 0.145, Kurtosis = −0.280). On mastery and performance-avoidance, this study found (M = 3.025, SD = 0.538) and normality test result (Skewness = 0.090, Kurtosis = −0.312). The reliability test results were shown using the composite reliability (CR) and Cronbach's Alpha coefficient. According to Gravetter et al. (2021), an internal consistency coefficient of>.60 was acceptable, while Urbina (2014) reported that a Cronbach's Alpha coefficient of >.70 had good reliability. In this study, the coefficient α = 0.617 (mastery goal), 0.629 (performance goal), 0.749 (mastery avoidance and performance-avoidance), and AGQ-R was found to have good reliability. According to Hair (2019), the cut-off value for CR was greater than .70. Therefore, mastery (CR = 0.746) and performance (CR = 0.739) goals, as well as mastery-avoidance and performance-avoidance (CR = 0.869), showed good reliability.

3.1. Convergent Validity

Table 1 shows crucial evidence for the convergent validity of CFA model. Convergent validity assessed the degree to which a set of items, theoretically designed to measure a specific construct, were highly correlated and loaded together on the construct (Hair et al., 2019). This was empirically evaluated by examining AVE for each construct. Table 1 also presented evidence of convergent validity, assessed through AVE, which required values of ≥ .50 (Hair et al., 2019; Muthén & Muthén, 2017) to show that a construct accounted for more than half of the variance in its observed items. AVE values for "Mastery Goal" (0.721), "Performance Goal" (0.675), and "Mastery Avoidance and Performance-Avoidance" (0.778) all surpassed this threshold, providing strong evidence of convergent validity for each dimension. This result showed that the items in each dimension adequately converged to measure the respective underlying constructs.

3.2. Discriminant Validity

Item analysis

Item-Level Psychometric Properties (AGQ-R): Item bias, according to Bond & Fox (2018), was not the focus in item selection (Table 2). However, information about the existence of biased items was very influential on measurement accuracy. The accuracy of the measurement was impacted by bias (Urbina, 2014). When individuals with certain traits received preference in responding to questions compared to those with different traits, the item was considered biased. Differential item functioning (DIF) in the Rasch measurement model allowed identification of the item bias. It was advised that items marked with DIF ($p < .05$) be examined, and when required, amended or replaced. AGQ-R had DIF value above $p < .05$, as shown in Table 2.

Table 2 details the psychometric properties of AGQ-R at the item level (N=200), providing insights into item performance in each dimension. Factor loadings for all items across "Mastery goal," "Performance Goal," and "Mastery Avoidance and Performance-Avoidance" dimensions were generally strong, ranging from 0.43 to 0.79, suggesting that items adequately represent the intended constructs. The commonality (h2) values, ranging from 0.11 to 0.88, further

supported the proportion of variance in each item explained by its common factor. Item fit, assessed by Infit MnSq and Outfit MnSq, was significantly in acceptable ranges (generally between 0.5 and 1.5), suggesting that most items measured the respective constructs consistently without significant misfit. Item difficulty varied across dimensions, providing a spread of item locations on the latent trait continuum.

Table 2. Psychometric properties of AGQ-R at the item level (N = 220)

Dimension	Item	Factor loading	h^2	Infit MnSq	Outfit MnSq	Difficulty	DIF Contrast across Gender
Mastery goal	Item 1	0.60	0.58	1.08	1.07	-.025	0.50
	Item 3	0.79	0.40	1.50	1.43	0.30	0.53
	Item 7	0.52	0.42	1.14	1.29	0.93	0.72
Performance goal	Item 5	0.65	0.14	1.41	1.35	-0.24	0.55
	Item 9	0.62	0.41	0.88	0.86	-0.30	0.66
	Item 11	0.47	0.39	1.19	1.13	-0.21	0.32
Mastery Avoidance and Performance-Avoidance	Item 2	0.44	0.53	0.86	0.88	-0.30	0.66
	Item 4	0.54	0.64	0.81	0.90	-0.05	0.73
	Item 8	0.63	0.11	0.38	0.90	-0.05	0.59
	Item 6	0.57	0.31	0.47	0.83	-0.47	0.11
	Item 10	0.60	0.88	0.90	0.92	0.05	0.54
	Item 12	0.43	0.40	1.40	0.41	0.38	0.39

Note: MnSq mean square error, DIF differential item functioning

*Extraction method: Oblimin rotation with Kaiser normalization. h2 = commonalities
cDIF contrast > .5 represents substantial DIF
dDIF contrast across gender = difficulty for males-difficulty for females

A critical result from the item-level analysis concerned DIF across gender. Several items exhibited a cDIF contrast above 0.50, suggesting substantial DIF. There was significant bias in Items 7 (0.72), 9 (0.66), and 4 (0.73) in "Mastery Goal," "Performance Goal," and "Mastery Avoidance and Performance-Avoidance," respectively. This result showed that male and female participants with the same underlying trait level responded differently to the items, suggesting a potential area for future item refinement to ensure equitable measurement across genders.

3.3. Validity Testing Results

Validity testing results for the multidimentional model of AGQ-R showed that all 12 AGQ-R items were valid in measuring the achievement goals construct. All items exhibit strong factor loadings, ranging from 0.43 to 0.79, consistently meeting or exceeding the common acceptance thresholds of 0.40 or 0.50. The analysis showed that all these factor loadings were statistically significant, with Est/SE (equivalent to t-values) well above 1.96 and p-values of 0.000 or 0.001. This result suggested that the relationships between the items and the latent construct were robust and not due to chance. Specifically, items measuring Mastery Goals consistently showed strong loadings, while those related to Performance Goals had adequate validity. Consequently, these data provided strong evidence that the 12 AGQ-R items used in this study effectively and reliably measure various dimensions of achievement goals in the investigated student population.

CFA results showed that the instrument exhibited a strong factorial structure, along with sufficient reliability and validity. This substantiated its utility for accurately and consistently assessing the intended dimensions. After performing CFA, a critical subsequent step was to review the goodness-of-fit indicators. This review was essential to ascertain the adequacy of the proposed theoretical model, as the indicators offered an evaluation of how closely the observed data conform to the specified theoretical structure. Scrutinizing the goodness-of-fit was significant for validating both the hypothesized factorial structure and the internal consistency of the instrument. (Gravetter et al., 2021). Table 4 shows the goodness-of-fit indices of CFA model.

Table 3 is a summary of the multidimentional model for AGQ-R. The 12 AGQ-R items were valid in measuring achievement goals with a t-value range (9.779− 36.323) and a t-value ≥ 1.96 (Gravetter et al., 2021). The multidimentional model results showed that the model fit with $\chi 2$ (166) = 202.498, p = .0000, RMSEA = 0.045, 90% CI = 0.051− 4.493, probability RMSEA = 2.227, CFI =1.000, TLI = 1.000, SRMR = 0.046. Urbina (2014) reported that the RMSEA coefficient ranged from .03 to .08. Xia and Yang (2018) stated that RMSEA > .80 as a reasonably fit. This

study reported RMSEA value of 0.045, and SRMR was found to be .056 (below .08). Therefore, the absolute fit measure has achieved goodness of fit. For incremental fit indices, CFI and TLI were in the range of 0−1, with higher coefficients, a more fit (Hair et al., 2019). Figure 1 shows the multidimentional diagram of AGQ-R model, with a factor loading of 12 items, ranging from 0.665 to 1.000. According to Hair et al. (2019), factor loading must be at a minimum of .50 and ideally 1.00. Therefore, 12 items were declared valid in measuring achievement goals.

Table 3. Multidimentional model of AGQ-R (N=220)

	Item	Factor loading	Est/SE	P-value
Mastery goal	Item 1	0.60	17.978	0.001
	Item 3	0.79	9.991	0.003
	Item 7	0.52	11.678	0.000
	Item 5	0.65	15.451	0.000
Performance goal	Item 9	0.62	10.876	0.000
	Item 11	0.47	13.456	0.000
	Item 2	0.44	19.001	0.000
Mastery Avoidance and Performance-Avoidance	Item 4	0.54	12.101	0.000
	Item 8	0.63	14.178	0.000
	Item 6	0.57	11.008	0.000
	Item 10	0.60	16.001	0.000
	Item 12	0.43	12.105	0.000

Note: Est/SE (T-value) = >1.96; *p-value*= >0.005

Table 4. Goodness-of-fit (GOF) indices of CFA model

Goodness-of-fit Index	Fit Index	Model Index	Model Evaluation
Chi-square $\chi 2$	>0.05	202.498	Model fit criteria
Koefisien RMSEA	<0.05	0.045	Model fit criteria
90 percent C.I	<0.05	0.051 – 4.493	Model fit criteria
Probability RMSEA	>0.05	2.227	Model fit criteria
CFI	1- > 0.80	1.000	Model fit criteria
TLI	1- > 0.80	1.000	Model fit criteria
SRMR	<0.05	0.046	Model fit criteria

Note: CFI; Comparative Fit Index; TLI; Tucker-Lewis Index; Standardized Root Mean Square Residual

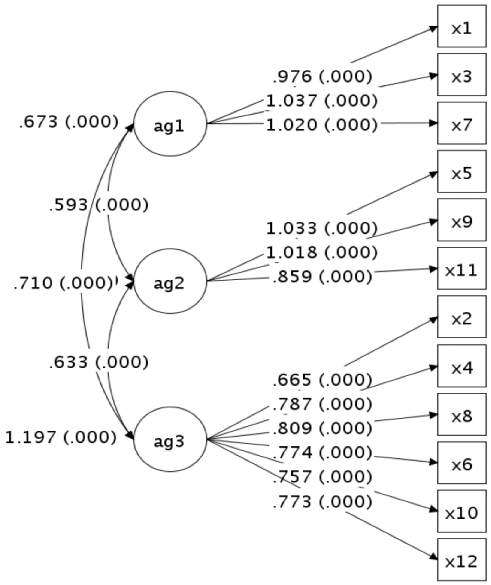

Figure 1. Multidimensional Model of AGQ-R (n = 220)

CFA results showed that the instrument possessed a robust factorial structure, complemented by adequate reliability

and validity, supporting its use for accurately and consistently measuring the proposed dimensions. After performing CFA, a critical subsequent step included reviewing the goodness-of-fit indicators to ascertain the adequacy of the proposed theoretical model. These indicators offered an evaluation of how closely the observed data conformed to the specified theoretical structure. The scrutiny of goodness-of-fit was important for validating both the hypothesized factorial structure and the internal consistency of the instrument. Key goodness-of-fit indicators evaluated were RMSEA ≤.08, SRMR ≤.08, CFI ≥.95, and TLI ≥.90 (Hu & Bentler, 1999; Raykov & Marcoulides, 2011; Hair et al., 2019). Although the chi-square (χ^2) statistic was reported, its sensitivity to sample size meant it was considered less critical for the model fit evaluation.

3.4. Discussions

The CFA results revealed that the instrument had a strong factorial structure, as well as adequate reliability and validity, indicating that it could be used to accurately and consistently measure the proposed dimensions. Items measuring Mastery Goals had consistently high loadings, whereas those related to Performance Goals had adequate validity. As a result, these findings provided strong evidence that the 12 AGQ-R items used in this study effectively and reliably measured various aspects of achievement goals in the studied student population. A key finding from the item-level analysis concerned DIF across gender. Items in the "Mastery Goal," "Performance Goal," and "Mastery Avoidance and Performance-Avoidance" categories all showed significant bias. This result showed that male and female participants with the same underlying trait level responded differently to the items, suggesting a potential area for future item refinement to ensure equitable measurement across genders.

This result, which favored a three-factor structure over the proposed four-factor model, was consistent with a growing body of literature suggesting that the classic 2×2 framework might not have been applicable across all populations and contexts (Chowning & Campbell, 2009; Elliot & Murayama, 2008; Tomczak et al., 2024; Yeh et al., 2019). Additionally, compared to the 2x2 framework models, Simamora & Mutiarawati (2021) found that the trichotomous model was the most effective model for first-year undergraduate students. AGQ-R scores fit a three-factor model that consolidated avoidance into a single domain (Cook et al., 2017). The development of a robust three-factor model in the current sample of undergraduate students provides crucial evidence for the cultural and contextual nuances in how achievement goals are manifested and perceived. Specifically, this shows that for the particular student population, the distinction between certain goal orientations, particularly in the avoidance dimension, might be less salient or coalesce into broader constructs. The results of this study provide empirical support for separating relatively low agreement variance towards goal and performance goals to develop a measurement of goal reflectivity in students. Future studies on AGT can take a variety of goals, not limited to undergraduate students. Future research could reach a wider range of participants, considering that Indonesia consists of many provinces and cities. To obtain more representative results, a probability sampling technique could be used. Psychometric analysis was conducted by examining internal consistency. Future research could conduct validity tests by examining associations with other variables.These include longitudinal studies to examine stability over time or studies exploring its utility in different educational settings beyond the university.

4. Conclusion

This study aimed to explore the psychometric properties, including confirmatory factor and invariance analysis, of the AGQ-R scale, to determine the best fit for identifying undergraduate students' achievement goals. Elliot & Murayama (2008) developed AGQ-R to measure beliefs about mastery-goal, mastery-avoidance, performance-goal, and performance-avoidance goals. Based on the result, the three-factor model is the most appropriate representation and a valid and reliable instrument to measure achievement goals. These results provide empirical support for separating relatively low agreement variance towards goal and performance goals to develop a measurement of goal reflectivity in students. This validated psychometric instrument specifically tailored for the Indonesian higher education context to assess student motivation and design culturally-aligned interventions. These insights allow for the development of teaching strategies that better support student engagement and academic success in diverse, collectivist learning environments. Future studies on AGT can take a variety of goals, not limited to undergraduate students

Acknowledgements

The authors would like to express their sincere gratitude to all participants who generously contributed their time and

effort to this study. Their cooperation and willingness to participate were essential to the successful completion of this research.

References

Apostolou, M. (2013). Psychometric Properties of the Achievement Goal Questionnaire-Revised Administered to Greek University Students. *Psychological Reports*, *113*(2), 380–403. https://doi.org/10.2466/14.11.PR0.113x22z1.

Becerra-Posada, T., García-Montes, P., Sagre-Barbosa, A., Carcamo-Espitia, M. I., & Herazo-Rivera, J. D. (2022). Project-based learning: The promotion of communicative competence and self-confidence at a State High School in Colombia. *How, 29(2), 13-31*.

Boiché, J. C. S., Sarrazin, P. G., Grouzet, F. M. E., Pelletier, L. G., & Chanal, J. P. (2008). Students' motivational profiles and achievement outcomes in physical education: A self-determination perspective. *Journal of Educational Psychology, 100*(3), 688–701. https://doi.org/10.1037/0022-0663.100.3.688

Bond, T. G., & Fox, C. M. (2018). *Applying the Rasch model: Fundamental measurement in the human sciences*. Psychology Press. https://www.taylorfrancis.com/books/mono/10.4324/9781410614575/applying-rasch-model-trevor-bond-christine-fox

Campbell, H. L., Barry, C. L., Joe, J. N., & Finney, S. J. (2008) Configural, metric, and scalar invariance of the modified achievement goal questionnaire across African American and white university students. *Educational and Psychological Measurement, 68(6), 988-1007.* https://doi.org/10.1177/0013164408318766

Chazan, D. J., Pelletier, G. N., & Daniels, L. M. (2022). Achievement goal theory review: An application to school psychology. *Canadian Journal of School Psychology, 37(1), 40-56*. https://doi.org/10.1177/08295735211058319

Cheng, X. (2023). Looking through goal theories in language learning: A review on goal setting and achievement goal theory. *Frontiers in psychology*, *13*, 1035223. https://doi.org/10.3389/fpsyg.2022.1035223

Chowning, K., & Campbell, N. J. (2009). Development and validation of a measure of academic entitlement: Individual differences in students' externalized responsibility and entitled expectations. *Journal of Educational Psychology*, *101*(4), 982–997. https://doi.org/10.1037/a0016351

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.

Cook, D. A., Castillo, R. M., Gas, B., & Artino Jr, A. R. (2017). Measuring achievement goal motivation, mindsets and cognitive load: Validation of three instruments' scores. *Medical education, 51(10), 1061-1074. https://doi.org/10.1111/medu.13405*

Daniels, L. and Bulut, O. (2019). Students' perceived usefulness of computerized percentage-only vs. descriptive score reports: associations with motivation and grades. *Journal of Computer Assisted Learning, 36(2), 199-208*. https://doi.org/10.1111/jcal.12398

Deeba, F., & Ahmad, I. (2023). A Rasch Analysis of Achievement Goal Orientation Scale in Language Education Context. *NUST Journal of Social Sciences and Humanities*, *9*(2), 184-208. https://doi.org/10.51732/njssh.v9i2.171

Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist, 34(3), 169–189.* https://doi.org/10.1207/s15326985ep3403_3

Elliot, A. J., & McGregor, H. A. (2001). A 2× 2 Achievement Goal Framework. *Journal of personality and social psychology, 80(3), 501*. https://doi.org/10.1037/0022-3514.80.3.501

Elliot, A. J., & Murayama, K. (2008). On the measurement of achievement goals: Critique, illustration, and application. *Journal of Educational Psychology*, *100*(3), 613–628. https://doi.org/10.1037/0022-0663.100.3.613

Fokkema, M., Kuyper, H., & Van der Werf, G. (2024). Thriving and striving around the world: A cross-cultural examination of the relationship between achievement goals and flourishing. *Journal of Happiness Studies*, https://doi.org/10.1007/s10902-024-00828-6

Gravetter, F. J., Wallnau, L. B., Forzano, L.-A. B., & Witnauer, J. E. (2021). *Essentials of statistics for the behavioral sciences* (Edition 10). Cengage.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (Eighth edition). Cengage.

Hart, C., Mueller, C., Royal, K., & Jones, M. (2012). Achievement goal validation among african american high school students. *Journal of Psychoeducational Assessment, 31(3), 284-299.* https://doi.org/10.1177/0734282912466726

Herdian et al. 2021. Reliability and Correlation Item on Achievement Goal Orientation Scale of the Students College in Indonesia. *Nursing*, *2*, 0-5

Holzinger, K. J., & Swineford, F. (1937). The Bi-Factor Method. *Psychometrika*, *2*(1), 41–54. https://doi.org/10.1007/BF02287965

Kaplan, R. M. (2018). *Psychological testing: Principles, applications, & issues* (Ninth edition). Cengage Learning.

Kwon, H., Capraro, R. M., & Capraro, M. M. (2021). When I believe, I can: Success STEMs from my perceptions. *Canadian Journal of Science, Mathematics and Technology Education, 21(1), 67-85*. https://doi.org/10.1007/s42330-020-00132-4

Muthén, L. K., & Muthén, B. (2017). *Mplus user's guide: Statistical analysis with latent variables, user's guide*. Muthén & Muthén.

Muis, K. R., & Winne, P. H. (2012). Assessing the psychometric properties of the achievement goals questionnaire across task contexts. *Canadian Journal of Education/Revue canadienne de l'éducation*, *35*(2), 232-248

Ning, H. K. (2016). A psychometric evaluation of the achievement goal questionnaire–revised in singapore secondary students. *Journal of Psychoeducational Assessment, 35(4), 424-436.* https://doi.org/10.1177/0734282916629850.

Ratsameemonthon, L. (2017). The achievement goal questionnaire-revised for thai college students and asian context. *Electronic Journal of Research in Educational Psychology, 13(36), 369-386.* https://doi.org/10.14204/ejrep.36.14078

Rodliya, R. (2018). Reliabilitas dan validitas konstruk achievement goal questionnaire – revised (AGQR) versi Indonesia. [*Skripsi*]. Universitas Pendidikan Indonesia

Rosas, J. (2015). Validation of the achievement goal questionnaire – revised in argentinean university students (a-agq-r). *International Journal of Psychological Research, 8(1), 10-23*. https://doi.org/10.21500/20112084.64

Roberts, G. C., & Nerstad, C. G. (2020). Motivation: achievement goal theory in sport and physical activity. In The Routledge international encyclopedia of sport and exercise psychology (pp. 322-341).

Sánchez Rosas, J. (2015). Validation of the achievement goal questionnaire – revised in Argentinean university students (A-AGQ-R). *International Journal of Psychological Research*, *8*(1), 10–23. https://doi.org/10.21500/20112084.641.

Senko, C., Hulleman, C. S., & Harackiewicz, J. M. (2011). Achievement goal theory at the crossroads: Old controversies, current challenges, and new directions. *Educational Psychologist*, 46(1), 26–47. https://doi.org/10.1080/00461520.2011.538646

Simamora, B., & Mutiarawati, E. V. (2021). Achievement Goals Model Validation: Is the 2X2 Better than the Trichotomous?. *International Journal of Evaluation and Research in Education*, *10*(1), 142-149. https://doi.org/10.11591/ijere.v10i1.20869

Strunk, K. K. (2014). A factor analytic examination of the achievement goal questionnaire, revised: A three-factor model. *Psychological Reports, 115(2), 400-414. https://doi.org/10.2466/14.03.PR0.115c24z0*

Struck Jannini, A. V., Akdemir, Z., & Menekse, M. (2024). Achievement goal theory in STEM education: A systematic review. *Journal of Engineering Education*, *113*(4), 986-1007. https://doi.org/10.1002/jee.20585

Tomczak, M., Kleka, P., Tomczak-Łukaszewska, E., Bojkowski, Ł., & Walczak, M. (2024). Validation and personality conditionings of the 3 × 2 achievement goal model in sport. *Scientific Reports*, *14*, 1588. https://doi.org/10.1038/s41598-024-52075-7.

Tze, V. M., Li, J. C. H., & Daniels, L. M. (2022). Similarities and differences in social and emotional profiles among students in Canada, USA, China, and Singapore: PISA 2015. *Research Papers in Education*, *37*(4), 558-583. https://doi.org/10.1080/02671522.2020.1864760

Urbina, S. (2014). Essentials of Psychological Testing. In *Essentials of Psychology Testing Second Edition* (2nd ed.). John Wiley & Sons, Inc., Hoboken, New Jersey. https://doi.org/10.1002/9781394259458

Urhahne, D., & Wijnia, L. (2023). Theories of motivation in education: An integrative framework. *Educational Psychology Review*, *35*(2), 45. https://doi.org/10.1007/s10648-023-09767-9

Was, C. A., Al-Harthy, I., Stack-Oden, M., & Isaacson, R. M. (2017). Rango de identidad académica y su relación con la orientación de meta de logro. *Electronic Journal of Research in Education Psychology*, *7*(18), 627-652. https://doi.org/10.25115/ejrep.v7i18.1363

Williams, B., Onsman, A., & Brown, T. (2010). Exploratory Factor Analysis: A Five-Step Guide for Novices. *Australasian Journal of Paramedicine*, *8*, 1–13. https://doi.org/10.33151/ajp.8.3.93

Yeh, Y.-C., Kwok, O.-M., Chien, H.-Y., Sweany, N. W., Baek, E., & McIntosh, W. (2019). How College Students' Achievement Goal Orientations Predict Their Expected Online Learning Outcome: The Mediation Roles of Self-Regulated Learning Strategies and Supportive Online Learning Behaviors. *Online Learning*, *23*(4), Article 4. https://doi.org/10.24059/olj.v23i4.2076