Journal of Ecohumanism
2025
Volume: 4, No: 1, pp. 2001 – 2016
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i1.6020

Using the Graded Response Model to Validate and Standardize the Academic Self Concept Scale for University Students in Saudi Arabia

Hesham Yahya Ali Aljubaily[1]

Abstract

This study aimed to standardize the Academic Self Concept (ASC) Scale for use with Saudi university students. The focus was on evaluating its psychometric properties, including reliability, validity, and ability to measure latent traits using Item Response Theory (IRT), specifically the Graded Response Model (GRM).Data were collected from 894 university students using the ASC scale, comprising 40 items across four domains: Mathematics, Verbal, Academic, and Problem-Solving. IRT-GRM was applied to assess item difficulty, discrimination, and test information. Structural Equation Modeling (SEM) was used to confirm the scale's construct validity, while measures marginal reliability was used to measure the scale's reliability.IRT analysis revealed strong psychometric properties, with high item discrimination (0.40–1.68), good model fit indices, and high test reliability (marginal reliability = 0.93). The ASC scale demonstrated excellent psychometric properties. Construct validity was confirmed through SEM, SEM confirmed the four-factor structure, with strong fit indices (CFI = 0.96, TLI = 0.95, RMSEA = 0.07). Composite Reliability (CR), for the four domains were ranging from 0.85 to 0.92, served as evidence of convergent validity. Discriminant validity was achieved with values of HTMT < 0.85. The ASC scale demonstrated excellent psychometric qualities, supporting its use in measuring academic self-concept among Saudi students. Recommendations include its application in higher education for academic support programs and further validation in diverse populations. Future research should explore cultural influences and develop digital assessment tools.

Keywords: *Academic Self Concept, Item Response Theory, Graded Response Model, Structure Equation Model, Marginal Reliability.*

Introduction

Academic self-concept is considered a highly important field of study in educational psychology, as it includes the student's perception of their academic abilities and competence. It plays a crucial role in shaping educational experiences, influencing motivation, behavior, and academic performance. Academic self-concept is multidimensional and reflects beliefs about abilities and various skills in academic fields. Therefore, understanding and measuring academic self-concept is essential for educators and researchers who aim to enhance student outcomes and design educational programs.

Academic self-concept has been identified as an important indicator of academic achievement, as research suggests that a positive academic self-concept is associated with academic achievement. Students with a positive academic self-concept tend to achieve higher academic performance, have greater motivation, and hold a more positive attitude towards learning (Marsh & Martin, 2011). Marsh et al. (2018) observed that students who have a strong belief in their academic abilities are more likely to engage in challenging tasks and persevere in the face of difficulties. This positive correlation is often attributed to the fact that students who perceive themselves as competent are more engaged in academic tasks, exert effort, and persevere in the face of challenges. Conversely, students with a low academic self-concept may suffer from a lack of motivation, withdrawal, and ultimately weaker academic outcomes (Eccles & Wigfield, 2002).

The multidimensional nature of academic self-concept requires a precise approach to its measurement. The Self-Description Questionnaire (SDQIII), developed by Marsh (1984), is one of the most comprehensive tools for measuring academic self-concept across different domains. It includes domains such as mathematics, verbal skills, academic ability, and problem-solving; allowing for a detailed evaluation of the student's self-perceptions in these areas. Distinguishing between these domains is crucial, as students often have distinct self-concepts in different subjects. For example, a student may have a high self-concept in

[1] Department of Psychology, College of Social Sciences, Al Imam Mohammad Ibn Saud Islamic University, Riyadh, Saudi Arabia, Email: hyaljubaily@imamu.edu.sa, (Corresponding Author), https://orcid.org/0009-0009-5641-5217

Journal of Ecohumanism
2025
Volume: 4, No: 1, pp. 2001 – 2016
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i1.6020

mathematics but a lower self-concept in verbal skills; highlighting the importance of domain-specific measurement. Therefore, Item Response Theory (Item Response Theory) offers many advantages compared to Classical Test Theory (Classical Test Theory) in evaluating psychometric properties. It provides detailed information about the characteristics of each item, including indicators of difficulty and discrimination, allowing for more accurate measurement of the underlying construct of the measured concept, and better detection of item functions across different levels of the latent trait (Embretson & Reise, 2000). The standardization of the Academic Self-Concept Scale in the Kingdom of Saudi Arabia is of particular importance due to the region's unique cultural and educational context. Therefore, this study aims to contribute to providing a reliable tool for evaluating the academic self-concept of Saudi university students by using the Graded Response Model (Graded Response Model) of Item Response Theory to assess its psychometric properties.

Theoretical Framework and Previous Studies

The academic self-concept is based on the broader framework of self-concept theory, which assumes that individuals develop perceptions of themselves based on their experiences and interactions in different domains(Marsh & O'Neill, 1984). Specifically, the academic self-concept is a distinct aspect of the self-concept domain, focusing on an individual's self-assessment in academic environments. It differs from the general self-concept, which encompasses a broader range of self-perceptions outside academic contexts(Marsh & Martin, 2011). The distinction between general and academic self-concept is crucial, as individuals may possess varying levels of self-concept across different domains. According to Marsh's model (1990), the academic self-concept is influenced by actual academic achievement, social comparisons with peers, and feedback from teachers and parents. This multifaceted nature of self-concept indicates that students' beliefs about their academic abilities are shaped by intrinsic and extrinsic factors. The theoretical foundation of the academic self-concept is based on the multidimensional and hierarchical model proposed by Marsh, which suggests that the self-concept is structured in a way that reflects. The self-concept consists of multiple aspects, each related to different aspects of an individual's life, with the general self-concept at the top, and specific self-concepts, such as the academic self-concept, forming the base (Brunner, et al.,2010). Marsh (Marsh & O'Neill,1984) introduced the Self-Description Questionnaire III (SDQIII), a widely used tool that includes the academic self-concept, which encompasses many sub-domains. The academic self-concept section of the scale includes four domains: abilities in mathematics, verbal skills, general academic abilities, and problem-solving skills. These domains are crucial because students often differentiate their self-concept based on specific academic subjects (Alkhateeb, et al.,2022). Understanding these domains of the academic self-concept in general is vital for developing programs designed to enhance students' academic self-concept in various educational fields.

Many studies have developed and standardized scales to measure students' academic self-concept using various psychometric methods. For example, the study by Matovu (Matovu,2014) aimed to standardize the Academic Self-Concept Scale among university students in Malaysia. The study was conducted on 280 male and female students from various colleges and academic levels. Structural Equation Modeling was used to evaluate the impact of academic self-concept on academic achievement. The study also employed confirmatory factor analysis to assess the psychometric properties of the Academic Self-Concept Scale. The outcomes indicated the appropriateness of fit indicators, such as the chi-square for the model. Convergent validity was supported by acceptable values of average variance extracted for the dimensions of academic confidence and academic effort, while differential validity was confirmed, as the values of average variance extracted were greater than the squared correlation. The study concluded that the scale is effective for measuring academic self-concept in diverse university categories. In a study conducted by Granero-Gallegos et al. (Granero-Gallegos et al., 2021), the aim was to standardize the Spanish version of the Academic Self-Concept Scale. The study included 681 participants from undergraduate and master's students. The study employed confirmatory factor analysis and exploratory structural equation modeling to evaluate the Academic Self-Confidence Scale, showing excellent model fit indicators ($\chi 2 = 2.45$, CFI = 0.975, RMSEA = 0.059). The reliability results indicated acceptable values for the two factors (0.76 for academic confidence, 0.78 for academic effort), with acceptable average variance extracted values, which support convergent validity, while differential validity was established through the correlation between the

Journal of Ecohumanism
2025
Volume: 4, No: 1, pp. 2001 – 2016
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i1.6020

two factors. It was concluded that the Academic Self-Concept Scale is a valuable tool for enhancing academic self-concept and engagement among university students. In the Philippines, Lirio et al. (Lirio et al., 2022) conducted a study to investigate the structural factors of the Academic Self-Concept Scale. The study involved a sample of 94 undergraduate students. Exploratory factor analysis and structural equation modeling were used, leading to the identification of a structure consisting of two factors: positive academic self-concept (confidence) and negative academic self-concept (self-doubt). The model's validity was verified using model fit indicators, which included indicators ($\chi 2$, GFI, CFI, RMSEA, TLI); indicating acceptable fit, and the scale's validity was verified through factorial validity, which indicates the presence of two factors for the scale. (Esnaola et al., 2023) conducted a study to compare the academic self-concept across cultures in Spain and China. A sample of 651 adolescents was analyzed, and the Short Form of the Self-Description Questionnaire II (SDQII-S) was administered to them. The researchers applied Structural Equation Modeling and exploratory graph analysis to examine the internal structure of academic self-concept and investigate measurement invariance across countries. The psychometric analyses confirmed the multidimensional nature of academic self-concept, according to adequate relevance indicators (CFI, RMSEA). The results of the high confirmatory analysis and structural equation modeling indicated that the three-factor model was a better fit for both sub-samples, with significant factor loadings ranging from .525 to .939. The reliability coefficients, Cronbach's Alpha, and Omega coefficient were good among the Spanish and Chinese sub-samples. Notably, the results of the network psychometric measurement showed cultural differences in the relationships between verbal and mathematical factors, where the school dimension was more associated with verbal factors among Spanish students, and with mathematical dimensions among Chinese students.

Study Problem

The academic self-concept forms an integral part of understanding educational experiences and the outcomes achieved by students. The academic self-concept is considered a highly important psychological construct that reflects students' perceptions of their academic abilities and influences their motivation, engagement, and overall academic performance (Al-Forti et al., 2021). Research conducted on academic self-concept has demonstrated its significant and lasting impact on academic outcomes. For example, Marsh and O'Mara (Marsh & O'Mara, 2008) found that academic self-concept is a strong predictor of academic achievement, and that students with a high academic self-concept are more engaged, perform better in academic tasks, show greater perseverance, and achieve higher grades. This relationship is also supported by the findings of Al-Maliki (2021) and Mahmoud (2017), which indicate a positive correlation between academic self-concept, academic performance, and academic engagement among university students in the Kingdom of Saudi Arabia. Their study emphasized the importance of fostering an educational environment that supports students' self-beliefs, ultimately enabling them to realize their potential and excel academically. Alkhateeb and others (Alkhateeb, et al., 2022) found a significant positive relationship between academic self-concept and academic achievement among university students in Qatar. Dogan and Durmus also found a correlation between academic self-concept and academic self-efficacy among university students in Turkey, indicating that students with a higher academic self-concept also reported greater self-efficacy, which in turn positively enhances academic performance (Dogan & Durmus, 2021).

Despite the importance of academic self-concept in the educational process, there is a lack of standardized, culturally relevant measures that are applicable to Saudi university students. The majority of current studies on academic self-concept have been conducted in Western countries, thus there is a need for tools that adapt and align with the culture to accurately measure concepts like academic self-concept in different environments (Fernández et al., 2021). Cultural factors such as language, educational systems, and societal norms significantly affect students' academic abilities. Despite its comprehensiveness and wide applicability in Western contexts, it may not be directly applicable to non-Western cultures without appropriate standardization and validation. The unique educational, cultural, and social context of Saudi Arabia requires the development of a scale to ensure accurate and meaningful evaluation of students' academic self-concept. One of the fundamental challenges in standardizing the Academic Self-Concept Scale for Saudi university students is ensuring cultural relevance and linguistic accuracy. The translation of the scale is not merely a matter of linguistic conversion but also involves adapting phrases to align with cultural and educational

Journal of Ecohumanism
2025
Volume: 4, No: 1, pp. 2001 – 2016
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i1.6020

patterns. This process must consider differences in educational systems, teaching methods, and societal values that may influence how students perceive their academic self-concept. Without proper standardization, there is a risk of measurement bias, which can undermine the validity and reliability of the instrument (Van de Vijver & Tanzer, 2004).

Moreover, the use of Item Response Theory in the development and standardization of academic self-concept scales is limited, as many psychological and educational studies rely on Classical Test Theory, despite the modern theory's superiority in providing a more accurate analysis by evaluating items independently, thus offering a more detailed diagnosis of the scale (Meguellati et al., 2024). Traditional methods may not provide sufficient information about item function and the overall structure of the scale, whereas Item Response Theory offers a more advanced approach; allowing for detailed analysis of item characteristics, such as difficulty and discrimination parameters, and also provides consistent and accurate measurements across different levels of the latent trait (Glas,2008). The use of Item Response Theory, specifically the Graded Response Model in this study, is intended to provide a comprehensive evaluation of the standardized Academic Self-Concept Scale to ensure its suitability for the student community in Saudi universities. Therefore, this study aims to standardize the Academic Self-Concept Scale for university students in Saudi Arabia, thereby contributing to the provision of an important scale in the academic evaluation process.

Study Questions

- What are the psychometric properties of the Academic Self-Concept Scale when analyzed using the Graded Response Model for polytomous items based on the Item Response Theory among university students in Saudi Arabia?

- What is the reliability level of the Academic Self-Concept Scale among university students in Saudi Arabia?

- What are the validity indicators possessed by the Academic Self-Concept Scale among university students in Saudi Arabia?

Theoretical and Practical Significance of the Study

First: Theoretical Significance

This study aims to enrich the body of knowledge regarding academic self-concept, particularly in the academic environment of Saudi university students, by using Item Response Theory to standardize a scale for measuring academic self-concept. This research provides a methodological framework that enhances the psychometric properties of measuring academic self-concept and contributes scientifically to the cross-cultural understanding of academic self-concept and its implications for the academic teaching and learning process.

Secondly: Practical Importance

Providing a valuable practical tool for educators to apply to Saudi university students, and to evaluate and enhance students' self-beliefs regarding their academic abilities. Quantitative measurement of academic self-concept can help in providing targeted interventions aimed at improving students' academic self-concept, thereby preparing them for academic challenges and future personal and professional life, thus contributing to improving academic outcomes, enhancing educational practices, and student performance.

Study Objectives

Journal of Ecohumanism
2025
Volume: 4, No: 1, pp. 2001 – 2016
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i1.6020

The primary purpose of this study is to standardize the Academic Self-Concept Scale to suit the culture and community of Saudi university students. It also aims to verify the psychometric properties, and the reliability and validity indicators of the scale using the Graded Response Model of Item Response Theory.

Study Boundaries and Limitations

Objective Boundaries: The subject of this study is limited to standardizing the Academic Self-Concept Scale among university students, which includes the domains of mathematics, verbal skills, general academic ability, and problem solving.

Human Boundaries: The application was conducted on students from Al Imam Mohammad Ibn Saud Islamic University from various scientific colleges.

Spatial Boundaries: This study was conducted at Al Imam Mohammad Ibn Saud Islamic University in Riyadh, Saudi Arabia.

Temporal boundaries: The study was conducted during the first semester of the academic year 1446 AH.

Study Terms

Academic Self-Concept (Academic Self Concept)

"A form of self-concept related to the academic aspect, which includes academic confidence, academic effort, and academic self-regulation. It refers to the university student's view of self-respect, abilities, competence, skills, academic potential, and ability to address problems in ways that align with their capabilities and scientific and practical skills" (Al-Fortia et al., 2021, p. 308). The operational definition of academic self-concept for the current study is: university students' perceptions of their abilities and academic competencies, which were measured through responses to the Academic Self-Concept Scale, comprising four dimensions: mathematics, verbal, general academic ability, and problem-solving skills.

Graded Response Model (GRM)

"It is an extension of the two-parameter logistic model (for Item Response Theory) with binary response items, and it is used for items with multiple responses. It is suitable when responses to an item can be classified into more than two ordered responses, such as representing graded scores to evaluate problem-solving or ordered levels for agreement or frequency of a certain condition" (Auné et al., 2019, p. 50).

Methodology and Procedures

Study Methodology

The research design relies on the descriptive research method, specifically describing the psychometric properties of the standardized Academic Self-Concept Scale for university students through the use of the Graded Response Model of Item Response Theory, which included numerous psychometric analyses of tests and indicators to verify the validity, reliability, and the measured latent trait.

Study Population

The study population consists of students from Al Imam Mohammad Ibn Saud Islamic University, totaling 81,563 according to the university's website statistics (Al Imam Mohammad Ibn Saud Islamic University, n.d.).

Study Sample

Journal of Ecohumanism
2025
Volume: 4, No: 1, pp. 2001 – 2016
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i1.6020

Data was collected using a voluntary sample method, through responses to an electronic version of the Academic Self-Concept Scale, which was sent via email. The study sample included (894) university students, consisting of (379) male students and (515) female students. The study sample was diverse, encompassing various colleges and academic departments at the university, which were classified into theoretical and scientific colleges. The average academic average of the study sample was calculated, with a value of (4.2) out of 5. Table (1) illustrates the characteristics of the study sample.

Table (1). Characteristics of the Study Sample

Characteristics of the Sample	Categories	Number	Percentage
Gender	Male Student	379	42,4
	Female Student	515	57,6
	Total	894	100
Type of College	Theoretical Colleges	627	70,1
	Scientific Colleges	267	29,9
	Total	894	100

Study Instrument

The Academic Self-Concept Scale, derived from the Self-Description Questionnaire (SDQIII) developed by Marsh (1984), was used to evaluate academic self-concept. The scale includes four subscales: Mathematics, Verbal Skills, General Academic Abilities, and Problem-Solving Skills. Each subscale consists of 10 items, totaling 40 items for the scale. The scale also contains twenty reverse items. A high score on the scale indicates a high academic self-concept, while a low score indicates a low academic self-concept. The scale was translated into Arabic, and back-translation was performed by specialized translators. Each item is evaluated using a five-point Likert scale ranging from completely applicable (5 points) to not applicable (0 points).

Psychometric Properties of the Initial Version of the Scale

An initial version of the scale was applied to a pilot sample of (70) participants to verify the clarity of the scale items and the initial psychometric properties.

Firstly: Validity

Content Validity: The Academic Self-Concept Scale was evaluated by six experts specializing in psychology, measurement, and evaluation. The Lawshe (Lawshe,1975) method was used to calculate the content validity indicators for the scale items and the scale as a whole by measuring the level of agreement among the experts. The Content Validity Ratio (CVR) was calculated for each item of the scale, and the Content Validity Index (CVI) was also calculated for each item and for the scale as a whole. The value of (CVR) for each item was 1, and the value of (CVI) for each item and for the scale was 1, indicating that content validity was achieved. Some items were also revised and rewritten according to the reviewers' comments.

*Concurrent Validity***:** The concurrent validity of the scale was verified by calculating the correlation coefficient and the relationship between academic self-concept scores and students' cumulative GPA. The result of the statistical analysis showed a statistically significant positive correlation between academic self-concept and cumulative GPA, $r = .51$, $p < .01$. This means that higher academic self-concept scores are associated with higher cumulative GPAs, thus supporting the evidence of concurrent validity.

Secondly: Reliability

The results of verifying the reliability of the scale indicate that the Cronbach's Alpha value is 0.841. This result suggests that the scale has internal consistency and high reliability. The split-half method was also

Journal of Ecohumanism
2025
Volume: 4, No: 1, pp. 2001 – 2016
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i1.6020

used to evaluate internal consistency, with the Cronbach's Alpha value for the first part being 0.805 and for the second part being 0.726. These results indicate good and acceptable reliability for both halves.

Statistical Analysis

The Graded Response Model of Item Response Theory was used to analyze item parameters (difficulty, discrimination, and in and out fit) using the two-parameter logistic model (PL2), and marginal reliability was tested through the application of the program (IRTPRO). Additionally, the construct validity of the scale was verified by using Structural Equation Modeling (Structure Equation Model), which included conducting Confirmatory Factor Analysis using the program (JASP).

Study Results

Results of the first question of the study: " What are the psychometric properties of the Academic Self-Concept Scale when analyzed using the Graded Response Model for polytomous items based on the Item Response Theory among university students in Saudi Arabia?".

The psychometric properties of the scale were evaluated using the (GRM) model for polytomous items based on Item Response Theory. Table No. (2) illustrates the results of the analysis of Item Information Function (IIF) and Test Information Function (TIF), which include the ability level values of the items, the standard error (SE), the infit and outfit indicators of the model, and the discrimination and difficulty parameters for each item.

Table (2). Results of the Analysis of Item Information Function, and Discrimination and Difficulty Parameters

Item Information Function					Discrimination and Difficulty Parameters					
Item	Measure	S.E.	Infit	Outfit	a	$b1$	$b2$	$b3$	$b4$	$b5$
1	-0.2193	0.021	1.09	1.091	0.8	-3.1	-2.1	-1	0.3	1.49
2	-0.0072	0.021	1.18	1.179	0.5	-3.2	-1.6	-0.2	1.54	3.27
3	0.02473	0.021	1.08	1.081	0.5	-3.2	-1.8	-0.1	1.92	3.55
4	0.38691	0.023	1.15	1.145	0.9	-3.4	-2.4	-1.4	-0.5	0.56
5	-0.2074	0.021	0.97	0.972	0.7	-4.2	-2.4	-1.1	0.16	1.89
6	0.35212	0.022	1.03	1.031	1	-3.3	-2.3	-1.2	-0.4	0.89
7	-0.193	0.021	1.02	1.022	0.9	-3.2	-1.9	-0.8	0.27	1.33
8	0.38017	0.023	1.02	1.02	1.1	-3.1	-2.3	-1.2	-0.4	0.78
9	-0.0426	0.021	1.16	1.162	0.7	-2.1	-1.4	-0.5	0.84	2.1
10	0.22546	0.021	1.28	1.282	0.8	-2.3	-1.7	-1	-0	1.12
11	0.27327	0.022	1.2	1.197	0.9	-2.8	-1.9	-0.9	-0.1	0.81
12	-0.1899	0.021	0.99	0.992	0.9	-3.4	-1.9	-0.7	0.27	1.34
13	0.3812	0.023	0.95	0.951	1	-3.6	-2.6	-1.4	-0.3	0.9
14	-0.1332	0.021	1.02	1.024	0.8	-3	-1.8	-0.7	0.7	1.68
15	0.46149	0.024	1.03	1.033	1.1	-3.6	-2.4	-1.4	-0.6	0.37
16	-0.3581	0.023	0.85	0.848	0.9	-4.8	-3.2	-1.5	-0.2	1.09
17	0.03497	0.021	0.94	0.935	0.6	-3.3	-1.8	-0.3	1.15	3.28
18	-0.296	0.022	0.95	0.946	1.1	-3.4	-2.1	-1.2	-0	0.98
19	0.92434	0.033	1.33	1.328	1.2	-4.6	-3.4	-2.4	-1.8	-0.9
20	-0.5577	0.025	0.96	0.959	1.5	-3.3	-2.5	-1.7	-0.8	0.27

Journal of Ecohumanism
2025
Volume: 4, No: 1, pp. 2001 – 2016
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i1.6020

21	-0.2557	0.022	0.84	0.836	1.7	-2.6	-1.8	-0.9	0.07	0.94
22	-0.0546	0.021	1.1	1.095	1	-1.8	-0.9	0.1	0.96	2.05
23	-0.2247	0.021	0.77	0.774	1.7	-2.7	-1.7	-0.7	0.23	1.07
24	0.2457	0.022	0.82	0.816	1.4	-2.7	-1.7	-0.9	0.11	1.18
25	-0.4241	0.023	0.72	0.724	1.5	-3.4	-2.6	-1.6	-0.3	0.85
26	0.20187	0.021	1.06	1.064	1.4	-2.1	-1.3	-0.7	-0	1
27	-0.399	0.023	0.76	0.763	1.6	-3.3	-2.2	-1.3	-0.2	0.81
28	0.3411	0.022	0.98	0.98	1.6	-2.3	-1.7	-1	-0.3	0.68
29	-0.451	0.024	0.78	0.783	1.4	-3.8	-2.7	-1.5	-0.4	0.72
30	0.69697	0.028	1.29	1.287	1.5	-3.1	-2.5	-1.8	-1.2	-0.4
31	0.70618	0.028	1.15	1.148	1.2	-3.6	-2.8	-2	-1.1	-0.3
32	-0.3781	0.023	0.86	0.864	1.1	-3.9	-2.7	-1.4	-0.2	0.92
33	-0.061	0.021	1.11	1.109	0.4	-3.5	-1.8	0.45	2.17	4.31
34	-0.4915	0.024	0.94	0.938	1.2	-4	-2.6	-1.5	-0.5	0.45
35	0.51069	0.024	0.89	0.894	1.2	-4.3	-2.7	-1.6	-0.8	0.63
36	-0.6685	0.027	1.09	1.085	1.1	-4.7	-3.1	-2.1	-1.1	-0.1
37	0.12838	0.021	0.8	0.802	0.9	-3.6	-2	-0.6	0.73	2.36
38	-0.423	0.023	1.03	1.034	0.7	-5.5	-3.8	-2.1	-0.6	0.72
39	-0.2193	0.021	1.24	1.294	1	-3.4	-2.4	-1.4	-0.7	0.4
40	-0.2532	0.021	1.22	1.309	0.7	-4.2	-3.1	-1.6	-0.1	1.13

Table number (2) illustrates the measurement of items difficulty, where the value represents the position of each item on the latent trait scale (θ) being measured. The difficulty values of the items ranged from -0.66847 (item 36) to 0.92434 (item 19). The information on item performance, which resulted in negative values, indicates that it is easier to agree upon and that it is likely to be agreed upon as applicable to the sample participants. On the other hand, the items that resulted in positive values require higher levels of the trait to agree that these items apply to them. The table also illustrates the outcomes of the standard error of measurement, which indicates the accuracy of measuring the difficulty of the items. It is evident from the table that the standard error values are low, ranging from 0.0206 to 0.0328, indicating precise measurement estimates. The table also illustrates the relevance indicators for the items by presenting the infit and outfit indicators. The acceptable range for relevance extends between 0.5 and 1.5, with values closer to 1 indicating better relevance (de Ayala, 2022). The results showed that all items fall within the acceptable range of 0.5 to 1.5 for infit and outfit indicators, indicating acceptable relevance for all scale items.

Table number (2) also presents the discrimination coefficients and the differential thresholds (Thresholds) for the responses of each item of the scale. The estimates of the parameters of the graded model provide information about the discrimination among the items, (a) and the ability levels ($b5 - b1$) for each item according to the graded response model. The discrimination parameter was evaluated based on the criteria described by De Ayala and Baker (De Ayala, 2022; Baker, 2001) as follows: very low ($a < 0.35$), low to moderate ($0.35 \geq a < 0.65$), moderate ($0.65 \geq a < 1.34$), and high ($a \geq 1.34$). The values of the discrimination coefficients for the items ranged between 0.40 and 1.68, with no item being very low, and all items falling within the acceptable criteria range; indicating the ability of the items to discriminate between individuals at different levels of the latent trait. The ability parameters ($b5 - b1$) indicate the differential thresholds between responses on the item for the latent trait. Since there are six responses for each item, the number of differential thresholds is five. The ability parameters ranged from -5.54 to 4.31, with negative values indicating low endorsement of the item, making them items with a low level of ability, or easier items. Positive values indicate higher endorsement of the item, making them higher-level items

Journal of Ecohumanism
2025
Volume: 4, No: 1, pp. 2001 – 2016
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i1.6020

for endorsement or difficult items requiring a high academic self-concept. Most items showed appropriately distributed thresholds, covering low to high levels of the latent trait.

Test information Function (TIF) was examined to verify the overall accuracy of the scale across the spectrum of latent traits. The results in Table (3) showed high measurement accuracy within the range of latent traits (θ) from − 2.8 to 1.6, with (TIF) values exceeding the standard for good performance as described by Toland (Toland, 2013), which is the value of ten for (θ), and standard error values below 0.3 in most of the latent trait range. However, test function began to decline below 10 after the latent trait value (θ) = 2, and standard error values increased, indicating reduced measurement accuracy at levels near the upper limit of the trait.

Table (3). Test Information Function

Latent trait ((θ	Test information	Expected standard error
-2.8	15.07	0.26
-2.8	15.79	0.25
-2	16.15	0.25
-1.6	16.29	0.25
-1.2	16.3	0.25
-0.8	16.23	0.25
-0.4	16.04	0.25
0	15.67	0.25
0.4	15.02	0.26
0.8	13.96	0.27
1.2	12.35	0.28
1.6	10.33	0.31
2	8.26	0.35
2.4	6.47	0.39
2.8	5.07	0.44

The level of goodness-of-fit for each item of the scale was examined using the probability-based item fit indicators (S-χ2) for the multivariate response data. According to Toland (Toland, 2013), items with (p-values) less than 0.01 are considered unfit for the model, while items with (p-values) greater than 0.01 are considered fit for the model. The results showed that all items obtained (p-values) greater than 0.01 and were not statistically significant; which means that all the scale's items are suitable for the Item Response Theory model. Several indicators were also used to evaluate the overall fit of the scale to the Graded Response Model based on the Item Response Theory according to the criteria set by De Ayala (De Ayala, 2022). Table number (4) illustrates the outcomes of these indicators.

Table (4). Indicators of Scale Fit

				RMSEA 90% CI			
CFI	TLI	SRMR	RMSEA	Lower	Upper	AIC	BIC
0,784	0.770	0.0652	0.0676	0.0651	0.0701	88059	88627

The value of the root mean square error of approximation (RMSEA) was 0.0676, indicating an acceptable fit for the model used, as it falls within the range of 0.05–0.08. The confidence interval (0.0651–0.0701) supports this conclusion, as the upper limit does not exceed 0.08. The standardized root mean square residual (SRMR) was also used, with a value of 0.0652, indicating a good fit for the model. The Comparative Fit Index (CFI) was also used, with a value of 0.784, which does not meet the acceptable thresholds for model fit, indicating a lack of fit. The Tucker-Lewis Index (TLI) was also used, with a value of 0.770, indicating that the model does not show a good fit.

Journal of Ecohumanism
2025
Volume: 4, No: 1, pp. 2001 – 2016
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i1.6020

Results of the second research question: "What is the reliability level of the Academic Self-Concept Scale among university students in Saudi Arabia?".

The reliability of the Academic Self-Concept Scale was calculated using the marginal reliability method, and its value reached 0.93, indicating excellent reliability according to measurement standards (Baker & Kim, 2017). Marginal reliability was also calculated for each level of the latent trait through the Test Information Function (TIF) that was previously presented. Table (5) illustrates the marginal reliability values across the latent trait. The marginal reliability values for the latent trait levels ranged between 0.81 and 0.94. The high values of the marginal reliability indicate a high degree of accuracy in measuring the levels of latent traits.

Table (5). Marginal Reliability of Latent Trait Levels

Latent trait (θ)	Reliability
-2.8	0.93
-2.8	0.94
-2	0.94
-1.6	0.94
-1.2	0.94
-0.8	0.94
-0.4	0.94
0	0.94
0.4	0.93
0.8	0.93
1.2	0.92
1.6	0.90
2	0.88
2.4	0.85
2.8	0.81

Results of the third research question: "What are the validity indicators possessed by the Academic Self-Concept Scale among university students in Saudi Arabia?".

Structural Equation Modeling; (SEM) was used to verify the construct validity of the standardized Academic Self-Concept Scale, in order to determine the model's fit with the four domains of the scale. Several indicators were used to select the quality of fit and alignment with Structural Equation Modeling, and they were evaluated based on De Ayala's criteria (De Ayala, 2022). The following table illustrates the outcomes of the fit indicators for Structural Equation Modeling.

Table (6). Outcomes of the Fit Indicators for Structural Equation Modeling

Fit Indicator	Outcome	Model Fit
χ^2	p > .01	Inadequate
CFI	0,96	Good Fit
TLI	0,95	Good Fit
NFI	0,95	Good Fit
IFI	0,96	Good Fit
RMSEA	0,07	Acceptable Fit
SRMR	0,07	Good Fit
GFI	0,97	Good Fit

It is evident from Table (6) that all indicators were overall fit for Structural Equation Modeling according to the fit acceptance criteria, except for the Chi-square indicator, which may be due to its sensitivity to large sample sizes (Bergh,2015). Confirmatory factor analysis was used as a procedure within the analysis of structural equation modeling, which measures the factor loadings values of the scale items for the four

Journal of Ecohumanism
2025
Volume: 4, No: 1, pp. 2001 – 2016
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i1.6020

domains. Values equal to or greater than 0.40 are generally considered acceptable factor loadings (Tabachnick & Fidell, 2013). The results showed that all factor loadings for all domains ranged between 0.40 and 0.80, except for item number 33 (I wish I had more imagination and creativity). Excluding this item, all the remaining nine items for the problem solving domain range between 0.48 and 0.74, which are acceptable factor loading values. In general, the results of the confirmatory factor analysis indicate a suitable relationship between the scale items and the domains to which they belong. Consequently, these results indicate the construct validity of the academic self-concept scale for university students.

The convergent validity of the scale was verified by calculating the composite reliability values, Composite Reliability (CR), and the average variance extracted Average Variance Extracted (AVE) for the scale domains. The value of (AVE) should be 0.5 or higher for each dimension of the scale according to the standards, or the composite reliability value should be higher than the (AVE) value for each dimension and exceed 0.7 even if the (AVE) value is less than 0.5 (Jiang et al.,2022; Ma et al., 2022). Table (7) illustrates the outcomes of the convergent validity indicators and composite reliability for the four domains of the scale.

Table (7). Indicators of Convergent Validity and Composite Reliability for the Scale

Domains	AVE	CR
Mathematics	0,5	0,92
Verbal	0,4	0,88
Academic	0,5	0,90
Problem Solving	0,4	0,85

The results in Table (7) indicate the achievement of convergent validity for the domains of mathematics and academic, with AVE values for each equal to 0.5. However, they were less than 0.5 for the verbal and problem-solving domains. Nonetheless, all composite reliability values were higher than 0.7 and also higher than the average variance extracted values for all domains, which ranged between 0.85 and 0.92, thereby achieving convergent validity according to the previous criteria.

The Heterotrait–Monotrait (HTMT) method was used to verify the discriminant validity of the scale, ensuring that the correlation coefficient does not exceed the value of 0.85 as a criterion for discriminant validity (Hair et al., 2019). Table (8) illustrates the outcomes of the analysis of (HTMT).

Table (8). The analysis of Heterotrait–Monotrait (HTMT)

	Mathematics	Verbal	Academic	Problem Solving
Mathematics	1			
Verbal	0.105	1		
Academic	0.439	0.597	1	
Problem Solving	0.218	0.607	0.533	1

The results in Table (8) indicate that the correlations between all factors do not exceed a value of 0.85, and the values of the correlation coefficient between the domains ranged from 0.105 to 0.607. These results indicate the achievement of the discriminant validity of the scale. In general, the results of convergent and discriminant validity are also indicators of the achievement of the construct validity of the Academic Self-Concept Scale.

Discussion of Results

First: Discussion of the Results of the First Question

The items parameters (difficulty and discrimination), items' infit and outfit, and overall scale information were evaluated. The results of the item difficulty parameter (b) showed that individuals' abilities spanned a

Journal of Ecohumanism
2025
Volume: 4, No: 1, pp. 2001 – 2016
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i1.6020

wide range, indicating that the items effectively cover different levels of academic self-concept. The lower thresholds, starting from ($b1$), represent the easier or less agreeable items, while the higher threshold ($b5$) represents more difficult or more agreeable items. As for the discrimination coefficients, most items showed acceptable discrimination according to accepted standards. These values indicate the scale's ability to effectively discriminate between individuals at different levels of academic self-concept.

The results showed that the overall function of the scale is appropriate around the average of the latent trait, with high test information function values and low standard error (SE) values, indicating that the scale provides high accuracy and high information for individuals with latent trait levels near the average. This result is consistent with the principles of Item Response Theory, where the latent trait is typically centered around an average, providing maximum information and accuracy in this range (Zanon et al.,2016). However, it is noted that test accuracy decreases at higher levels of the latent trait, where the values of (TIF) decrease and the values of (SE) increase when the level of the latent trait exceeds a value of 2. This decrease in accuracy at the extreme ends is common when using item response theory models, necessitating the addition of items or a review of items that fall at the higher levels of the latent trait to ensure the scale provides sufficient accuracy across the entire range of the latent trait. In terms of the fit of the items and the scale as a whole, the outcomes showed good fit for the items and the scale. The fit indicators for the items using (S-χ^2) analysis indicated that all items fit the (GRM) model well. Similarly, the results of the model fit indicators for the scale as a whole indicated a good fit, as the values of (RMSEA) and (SRMR) were within acceptable limits, suggesting a good fit for the model. Despite the low values of (CFI) and (TLI) for model fit, other results support the fit of the scale model used. Many recent studies support that (RMSEA) and (SRMR) are key indicators for assessing model fit, especially for ordinal or polytomous models, such as the Graded Response Model, due to their ability to provide a more accurate evaluation of overall and absolute model fit compared to indicators like (CFI) and (TLI), which are more sensitive to sample size and model complexity (Shi et al., 2019; Pavlov et al., 2021; Shi et al., 2020). Many studies support the importance of relevance indicators for evaluating model acceptance. However, relying solely on specific indicators can overlook other important outcomes and does not provide a complete picture of model fit. Model fit evaluation should include a comprehensive assessment using various methods of relevance indicators to achieve a thorough understanding of the scale's effectiveness and its compatibility with the model used (McNeish & Wolf, 2021; Schermelleh-Engel et al., 2003; Goretzko et al., 2024).

The results of the current study align with studies (Lirio et al., 2022; Matovu, 2014; Granero-Gallegos et al., 2021; Esnaola et al., 2023;) in terms of the suitability of the scale for model fit indicators, as the Academic Self-Concept Scale demonstrated very suitable psychometric properties when standardized in different cultures. However, each study used different psychometric methods and model fits to verify the scale structure. The current study differed from the study by Lirio et al. (Lirio et al, 2022), which relied on exploratory factor analysis and structural equation modeling to analyze their two-factor model: positive and negative. Matovu (Matovu, 2014) and Granero-Gallegos et al. (Granero-Gallegos et al., 2021) focused on using confirmatory factor analysis and structural equation modeling to verify the scale structure. Esnaola et al. (Esnaola et al., 2023) used structural equation modeling and network psychometrics. The current study is distinguished by applying item response theory analysis using the graded response model and structural equation modeling.

The results showed that the standardized academic self-concept scale meets the acceptable psychometric criteria for the items and is considered suitable for application in the Saudi environment. The researcher believes that the psychometric results of the scale's standardization, in general, confirm its validity and reliability, and thus its use and application to university students. The good standardization results may be attributed to the nature of the scale, which reflects various academic and educational skills, such as reading, thinking, calculation, and academic perceptions, which university students possess at different levels in universities worldwide. This is also demonstrated by many previous studies (Marsh et al., 2002; Arens et al., 2020; Wang & Liou, 2007; Fernandez et al., 2021) that have standardized (ASC) scales in different cultures, and their results also showed effectiveness, validity, and reliability of the scale in different countries.

Secondly: Discussion of the Results of the Second Question

Journal of Ecohumanism
2025
Volume: 4, No: 1, pp. 2001 – 2016
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i1.6020

The results of this study demonstrate the high reliability of the Academic Self-Concept Scale through the use of marginal reliability consistent with the Item Response Theory model according to psychometric standards. The reliability result indicates the consistency and accuracy of the scale for evaluating the academic self-concept of university students in the Saudi environment. The marginal reliability method was used because it is an appropriate method, even derived from Item Response Theory, as it takes into account the full range of the latent trait and provides an overall measure of the scale's accuracy by integrating the conditional standard errors (CSE) across all levels of the latent trait through test information function (Kılıç et al., 2023; Toland, 2013). The reliability result of the current study was similar to studies (Lirio et al, 2022; Matovu, 2014; Granero-Gallegos et al, 2021; Esnaola et al, 2023) in terms of having high reliability for academic self-concept scales. However, the methods of evaluating reliability varied across the studies, as most of these studies used classical test theory measures, such as Cronbach's Alpha, and the Omega coefficient. In the study by Lirio et al. (Lirio et al., 2022), reliability measurement was not among the study's objectives; therefore, it was not measured.

The researcher believes that the high reliability of the Academic Self-Concept Scale can be attributed to several procedural factors, such as the good translation of the scale, the clarity of the items, the modification of some items to fit the Saudi cuture according to expert opinion, and the examination of the initial psychometric properties of the scale through the pilot study. Such factors lead to enhanced reliability accuracy and reduce measurement error (DeVellis, 2016). Moreover, using the ideal method for calculating reliability, which is fully compatible with Item Response Theory, provides an accurate perception for revealing the scale's function across all levels. Kılıç et al. (2023), in their review of studies using the Item Response Theory model, found that most studies (84.21%) use traditional reliability estimates based on classical theory in studies that applied the Item Response Theory model. The use of marginal reliability is considered a better approach that suits Item Response Theory because it provides an accurate depiction of reliability across all levels of the latent trait and addresses the shortcomings of traditional measurement theory. This is because the information on item function takes into account the accuracy of measuring the full range of the latent trait, rather than providing a unified reliability estimate that does not measure the true construct of the scale (Auné et al., 2020; Toland, 2013).

Third: Discussion of the Results of the Third Question

In general, the overall results of the validity measurement showed that the standardized Academic Self-Concept Scale for university students in Saudi Arabia possesses construct, convergent, and discriminant validity. This was confirmed by the analyses of structural equation modeling, which indicated the scale's validity in measuring academic self-concept across its four sub-domains.

Confirmatory factor analysis was used, which is an analytical procedure within the various statistical analyses that fall under structural equation modeling. The outcomes of the confirmatory factor analysis showed that all items of the scale resulted in acceptable factor loadings with values higher than (0.4), except for one item (item 33). This item was retained to maintain the theoretical consistency of the scale and not to affect the overall validity outcome, which is considered good and consistent with other results. These results indicate a general alignment of the items with the structure of the latent trait intended to be measured. The results also showed the presence of convergent validity indicators for the scale, where the values of (AVE) were acceptable for the mathematics and academic domains, but not suitable for the verbal and problem-solving domains. However, the composite reliability values were high, exceeding the specified acceptance criterion, thus providing sufficient evidence of convergent validity, despite the low values of (AVE) in some sub-domains. A high composite reliability index is also used as another indicator to verify convergent validity (Jiang et al., 2022). The current study concluded that the Academic Self-Concept Scale possesses discriminant or differential validity. The results of the analysis of (HTMT) indicated that the correlations between factors were less than 0.85, which supports the discriminant validity of the scale. The low correlations between the domains in the current study confirm that the multidimensional structure of the scale is distinct, meaning that each of the four domains of the scale differentiates in measuring the domain it was designed to measure and contributes to measuring the trait of academic self-concept.

Journal of Ecohumanism
2025
Volume: 4, No: 1, pp. 2001 – 2016
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i1.6020

The current study agrees with studies (Matovu, 2014; Granero-Gallegos et al., 2021; Esnaola et al., 2023) in terms of using structural equation modeling to verify the construct validity of the Academic Self-Concept Scale, achieving fit indicators that align with structural equation modeling, thereby confirming the construct validity of the scale, and thus the results are similar in the validity of the scale across different cultural and educational contexts. However, the methods of evaluating validity were different in some studies. The current study differed in measuring construct validity from the Lirio et al. study (Lirio et al., 2022), as it relied on factorial validity through exploratory factor analysis. The current study was similar to the Matovu study (Matovu, 2014) in measuring convergent validity and discriminant validity, both of which were achieved with the scale. It was also similar to the Granero-Gallegos et al. study (Granero-Gallegos et al., 2021) in measuring only convergent validity, which was also achieved in that study. The current study is distinguished by using the method of analyzing (HTMT) to measure discriminant validity.

The slightly low values of (AVE) for two subdomains (verbal and problem-solving) may reflect the complexity of the construct composition of the academic self-concept trait, which consists of multiple and entirely different skills, such as the mathematics domain and the verbal domain. Recent studies highlight that academic self-concept relates to various skills, leading to differing internal dimensions across sub-domains. The study by Arens et al. (Arens, et al., 2020) reinforces this perspective, as their recent systematic review on the structure of academic self-concept highlighted the complexities of the academic self-concept structure and the dual multiplicity even within the dimensions of each academic domain. They believe that each domain of academic self-concept is divided into two dimensions: actual competence or skill in the domain, and opinion or attitude towards the domain. Therefore, this may lead to a decrease in the values of (AVE) for some sub-domains due to the precise and multifaceted nature of academic self-concept and its domains.

It is also possible for there to be a cultural influence on the domains of (verbal and problem solving) through educational and linguistic factors. For example, item (33): "I wish I had more imagination and creativity," may be interpreted differently among students and may hold different perspectives on creativity in Saudi academic environments compared to Western cultures. This is indicated by numerous studies that suggest specific cultural perspectives on creativity and problem-solving can influence students' responses on self-concept measures; which may affect how students perceive and respond to items related to specific skills in each domain (Lilla et al., 2021; Asbjørnsen et al., 2014; Smith et al., 2022). The results obtained by the study of Esnaola et al. (Esnaola et al., 2023) also support the cultural influence on academic self-concept, as their study found a significant difference in the relationship between verbal and mathematical factors across Spanish and Chinese cultures. For Spanish students, there was no significant correlation between the verbal domain and mathematics, whereas for Chinese students, there was a high correlation between the two domains. This result suggests that Chinese students may integrate their academic self-concepts across measurement domains more than Spanish students, and thus the dimensions of academic self-concept are interconnected differently in the two cultures.

Recommendations

The study recommends the following based on its results:

- Utilizing the scale in the field of student support, university counseling, guidance programs, and academic advising to identify the strengths and challenges students face in specific areas and skills, which helps in designing programs specifically to enhance students' academic self-concept, overall performance, and employability skills.

- The use of the scale in evaluating and developing academic programs, where the effectiveness of programs and their outcomes can be assessed, identifying areas where students exhibit a lower academic self-concept, and adjusting curricula or support mechanisms accordingly.

Suggestions

Journal of Ecohumanism
2025
Volume: 4, No: 1, pp. 2001 – 2016
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i1.6020

- Conducting longitudinal studies to explore the relationship between academic self-concept and various areas in the educational process, such as academic achievement, academic withdrawal, and career readiness.

- Conducting studies to explore the predictive validity of the scale concerning academic outcomes, such as dropout rates, academic average, and career readiness.

- Conducting studies in distance learning environments to assess academic self-concept in virtual educational settings.

References

Al-Fortia, S., Abu Mustafa, N. and Abu Mustafa, M. (2021). Predicting academic integration in light of academic self-concept and academic ambition among students of the Faculty of Education at Misurata University. Journal of Academic Research, 17, 325-303.

Al-Imam Muhammad bin Saud Islamic University. (N.D.). Facts and Statistics. www.imamu.edu.sa/about/Pages/statistics.

Alkhateeb, H. M., Abushihab, E. F., Alkhateeb, B. H., & Alkhateeb, R. H. (2022). Academic self-concept and its relationship to academic achievement among university students. International Journal on Social and Education Sciences (IJonSES), 4(4), 517-528.

Al-Maliki, p. (2021). Academic self-concept and its relationship to academic flexibility among male and female students of Taif University. Journal of the Faculty of Education - Menoufia University, (1), 142-193.

Arens, A. K., Yeung, A. S., Craven, R. G., & Hasselhorn, M. (2020). The structure of academic self-concept: A methodological review. Educational Psychologist, 55(3), 147-163.

Asbjørnsen, A. E., Obrzut, J. E., & Oyler, J. D. (2014). A cross-cultural comparison of verbal learning and memory functions in reading disabled American and Norwegian adolescents. Scandinavian journal of psychology, 55(2), 115–122. https://doi.org/10.1111/sjop.12108

Auné, S. E., Abal, F. J. P., & Attorresi, H. F. (2020). A psychometric analysis from the Item Response Theory (IRT) : Step-by-step modelling of a loneliness scale. Ciencias Psicológicas, 14(1), e2179. https://doi.org/10.22235/cp.v14i1.2179

Auné, S. E., Abal, F. J. P., & F Attorresi, H. (2019). Application of the Graded Response Model to a Scale of Empathic Behavior. International journal of psychological research, 12(1), 49–56. https://doi.org/10.21500/20112084.3753

Baker, F. B. (2001). The basics of Item Response Theory (IRT) .(2th ed.). ERIC Clearinghouse on Assessment and Evaluation.

Baker, F. B., & Kim, S.H. (2017). The Basics of Item Response Theory (IRT) Using R. Springer.

Bergh, D. (2015). Chi-squared test of fit and sample size: A comparison between a random sample approach and a chi-square value adjustment method. Journal of Applied Measurement, 16(2), 204–217.

Brunner, M., Keller, U., Dierendonck, C., Reichert, M., Ugen, S., Fischbach, A., & Martin, R. (2010). The structure of academic self-concepts revisited: The nested Marsh/Shavelson model. Journal of Educational Psychology, 102(4), 964–981.

De Ayala, R. J. (2022). The theory and application of Item Response Theory (IRT) (2th ed.). Guilford Press.

DeVellis, R. F. (2016). Scale development: Theory and applications (4th ed.). Sage Publications.

Dogan, I., & Durmus, G. (2021). Investigation of the relationship between academic self-concept and academic self-efficacy of university students receiving sports education. Journal of Educational Issues, 7(3), 68–84.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. Annual Review of Psychology, 53(1), 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153

Embretson, S. E., & Reise, S. P. (2000). Item Response Theory (IRT) for psychologists. Lawrence Erlbaum Associates.

Esnaola, I., Sesé, A., Azpiazu, L., & Wang, Y. (2023). Revisiting the academic self-concept transcultural measurement model: The case of Spain and China. British Journal of Educational Psychology, e12635. https://doi.org/10.1111/bjep.12635

Fernández, M., García, J., & López, M. (2021). Validation of the Academic Self-Concept Scale in the Spanish University Context. Psychology in Spain, 25(2), 150-165.

Glas, C. (2008). Item Response Theory (IRT) in educational assessment and evaluation. Mesure et évaluation en éducation, 31(2), 19–34. https:// doi.org/10.7202/1025005ar

Goretzko, D., Siemund, K., & Sterner, P. (2024). Evaluating Model Fit of Measurement Models in Confirmatory Factor Analysis. Educational and Psychological Measurement, 84(1), 123–144.

Granero-Gallegos, A., Baena-Extremera, A., Escaravajal, J. C., & Baños, R. (2021). Validation of the Academic Self-Concept Scale in the Spanish university context. Education Sciences, 11(10), 653. https://doi.org/10.3390/educsci11100653

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. European Business Review, 31, 2–24.

Jiang, Y. N., Wang, J. X., Chen, L. Y., Yao, J. J., Ni, L., Sheng, J. M., & Shen, X. (2022). Reliability and validity of the Composite Activity-Related Fall Risk Scale. Frontiers in Neurology, 13, 832691.

Jin, R., Wu, R., Xia, Y., & Zhao, M. (2023) What cultural values determine student self-efficacy? An empirical study for 42 countries and economies. Front. Psychol, 1- 14.

Journal of Ecohumanism
2025
Volume: 4, No: 1, pp. 2001 – 2016
ISSN: 2752-6798 (Print) | ISSN 2752-6801 (Online)
https://ecohumanism.co.uk/joe/ecohumanism
DOI: https://doi.org/10.62754/joe.v4i1.6020

Kılıç, A. F., Koyuncu, İ., & Uysal, İ. (2023). Scale development based on Item Response Theory (IRT) : A systematic review. International Journal of Psychology and Educational Studies, 10(1), 209-223. https://doi.org/10.52380/ijpes.202

Lawshe, C. H. (1975). The quantitative approach to content validity. Personnel Psychology, 28, 563-575.

Lilla, N., Thürer, S., Nieuwenboom, W., & Schüpbach, M. (2021). Exploring Academic Self-Concepts Depending on Acculturation Profile. Investigation of a Possible Factor for Immigrant Students' School Success. Education Sciences, 11(8), 432. https://doi.org/10.3390/educsci11080432

Lirio, G. A. C., Camba, R. L., Matibag, Y. B., & Conde, R. L. (2022). Academic self-concept of secondary science education students: A structural equation modeling approach. European Online Journal of Natural and Social Sciences, 11(1), 245 - 256.

Ma, L., Liang, Y., Guo, X., Wang, Y., & Xie, Q. B. (2022). Reliability and validity analysis of the Chinese version of Evaluation of Ankylosing Spondylitis Quality of Life (EASi-QoL). Clinical Rheumatology, 41(8), 2393–2401.

Mahmoud, H. (2017). Academic self-concept and level of academic ambition and their relationship to academic integration among a sample of university students.Journal of Educational Sciences, 25(2), 602-646.

Marsh, H. W. & Martin, A. J. (2011). Academic self-concept and academic achievement: Relations and causal ordering. British Journal of Educational Psychology, 81, 59-77.

Marsh, H. W. (1990). The structure of academic self-concept: The Marsh/Shavelson model. Journal of Educational Psychology, 82(4), 623–636. https://doi.org/10.1037/0022-0663.82.4.623

Marsh, H. W., & O'Mara, A. (2008). Reciprocal effects between academic self-concept, self-esteem, achievement, and attainment over seven adolescent years: Unidimensional and multidimensional perspectives of self-concept. Personality and Social Psychology Bulletin, 34(4), 542–552. https://doi.org/10.1177/0146167207312313

Marsh, H. W., & O'Neill, R. (1984). Self-Description questionnaire IIII: The construct validity of multidimensional self-concept ratings by late adolescents. In Journal of Educational Measurement, 21(2), 153–174. https://doi.org/10.1111/j.1745-3984.1984.tb00227.x

Marsh, H. W., Hau, K. T., & Grayson, D. (2002). Confirmatory factor analysis of multidimensional academic self-concept responses. Structural Equation Modeling, 9(2), 293–318. https://doi.org/10.1207/S15328007SEM0902_5

Marsh, H. W., Pekrun, R., Murayama, K., Arens, A. K., Parker, P. D., Guo, J., & Dicke, T. (2018). An integrated model of academic self-concept development: Academic self-concept, grades, test scores, and tracking over 6 years. Developmental Psychology, 54(2), 263–280. https://doi.org/10.1037/dev0000393

Matovu, M. (2014). A structural equation modelling of the Academic Self-Concept Scale. International Electronic Journal of Elementary Education, 6(2), 185-198.

McNeish, D., & Wolf, M. G. (2021). Dynamic fit index cutoffs for confirmatory factor analysis models. Psychological Methods, 28(1), 61-88. https://doi.org/10.1037/met0000425

Meguellati, S., Samia, A., Ferhat, A., Djelloul, A., & Ahmed Khalifa, Z. (2024). A critical analysis of the use of Classical Test Theory (CTT) in psychological testing: A comparison with Item Response Theory (IRT) (IRT). Pakistan Journal of Life and Social Sciences, 22(2), 9442-9449.

Pavlov, G., Maydeu-Olivares, A., & Shi, D. (2021). Using the standardized root mean squared residual (SRMR) to assess exact fit in structural equation models. Educational and Psychological Measurement, 81(1), 110–130. https://doi.org/10.1177/0013164420926231

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. Methods of Psychological Research, 8(2), 23–74.

Shi, D., Lee, T., & Maydeu-Olivares, A. (2019). Understanding the model size effect on SEM fit indices. Educational and Psychological Measurement, 79(2), 310–334.

Shi, D., Maydeu-Olivares, A., & Rosseel, Y. (2020). Assessing fit in ordinal factor analysis models: SRMR vs. RMSEA. Structural Equation Modeling: A Multidisciplinary Journal, 27, 1–15.

Smith, W., Hermida, J., & Güss, C. D. (2022). "Oh no, the forest is burning!" cultural differences in the complex problem-solving process only under high uncertainty. Frontiers in psychology, 13,1-18. https://doi.org/10.3389/fpsyg.2022.965623

Tabachnick, B. G., & Fidell, L. S. (2013). Using Multivariate Statistics (6th ed.). Boston, MA: Pearson.

Toland, M. D. (2013). Practical guide to conducting an Item Response Theory (IRT) analysis. The Journal of Early Adolescence, 34(1), 120–151. https://doi.org/10.1177/0272431613511332

Van de Vijver, F. J. R., & Tanzer, N. K. (2004). Bias and equivalence in cross-cultural assessment: An overview. Revue européenne de psychologie appliquée/European Review of Applied Psychology, 54(2), 119-135.

Wang, C. K. J., & Liou, M. H. (2007). The application of Item Response Theory (IRT) to assess the dimensionality and hierarchical structure of academic self-concept in Chinese students. Educational and Psychological Measurement, 67(5), 781-802. doi:10.1177/0013164406299103

Zanon, C., Hutz, C. S., Yoo, H., & Hambleton, R. K. (2016). An application of Item Response Theory (IRT) to psychological test development. Psicologia: Reflexão e Crítica, 29(1), 18. https://doi.org/10.1186/s41155-016-0040-x.