Pertanika J. Soc. Sci. & Hum. 33 (4): 1661 - 1681 (2025)

PERTANIKA JOURNALS

SOCIAL SCIENCES & HUMANITIES

Journal homepage: http://www.pertanika.upm.edu.my/

Examining the Psychometric Properties of the Westside Test Anxiety Scale Using the Rasch Model: Evidence from Nigerian Undergraduates

Abubakar Rabiu Uba[1], Ahmad Zamri Khairani[1]*, Muhammad Ibrahim Yakasai[2] and Abubakar Abdullahi[2]

[1]*School of Educational Studies, Universiti of Sains Malaysia, 11800 Penang, Malaysia*
[2]*Department of Education, Bayero University Kano, 700006 Kano, Nigeria*

ABSTRACT

The West Side Test Anxiety Scale (WTAS) is considered one of the most successful measures for determining which students have test anxiety deficits. However, the psychometric properties of this measure have not been adequately established, especially in Nigeria. Hence, the primary objective of the current study is to investigate the psychometric qualities of the 10-item WTAS in a sample of 300 undergraduates at a public university in Nigeria. This cross-sectional study employed a convenience sampling method. The investigation was conducted using the Rasch analysis measurement framework to investigate the following criteria: Rasch model's assumptions, construct validity, reliability evidence, Wright map, and differential item functioning (DIF). The results indicated satisfactory outcomes for most of the criteria, including DIF, construct validity, and reliability evidence. Specifically, the analysis revealed strong item reliability at 0.96 and acceptable person reliability at 0.70, indicating that the data met the expected standards by the Rasch model. However, it was discovered that the WTAS items were unable to target responders with higher abilities at the top of the assessed instrument on the Wright map. Consequently, we recommend that future revisions of the WTAS should include items specifically designed to target and differentiate among high-ability respondents. This adjustment would improve the scale's precision across a broader range of test-taker abilities.

Keywords: Nigeria, psychometric properties, Rasch model, undergraduates, westside test anxiety

ARTICLE INFO

Article history:
Received: 05 August 2024
Accepted: 14 March 2025
Published: 19 September 2025

DOI: https://doi.org/10.47836/pjssh.33.4.12

E-mail addresses:
abubakarbabura@student.usm.my (Abubakar Rabiu Uba)
ahmadzamri@usm.my (Ahmad Zamri Khairani)
Yakasaimi@gmail.com (Muhammad Ibrahim Yakasai)
a.abdullahi07@yahoo.com (Abubakar Abdullahi)
* Corresponding author

INTRODUCTION

Test anxiety is a negative psychological state that students experience in any kind of educational setting (Richard et al., 2023). It is normal to feel a little nervous before an exam due to human nature. On the other

e-ISSN: 2231-8534

hand, excessive and persistent anxiety is harmful and may affect a student's academic achievement (Muhammed et al., 2024). During exam times, anxiety is particularly prevalent due to pressure to perform well or worry about failure (Yuksel et al., 2018). Researchers argued that test anxiety may have adverse influence on students' performance, based on the type and intensity of the symptoms of anxiety (Hackel, et al. 2021; Pekrun & Stephens, 2012; Sanchez & Furlan, 2017; Von der Embse et al., 2018), student at the early stage of study are characterized by a lack of confidence, low self-efficacy, worry, and fear of failure, which in turn can influence their performance. In addition, several researchers have also discovered a negative correlation between all aspects of test anxiety and students' academic performance (Keith et al., 2019; Putwain et al., 2013; Raufelder & Ringeisen, 2016; Schnell et al., 2020).

Evidently, the ability to demonstrate skills and aptitudes on a test or set of exams may play a significant role in both academic advancement and entry into, or advancement within, a vocational sector (Khan et al., 2021). Given this, how people react to high exam anxiety may have a significant impact on educational settings, particularly for students in higher education who want to graduate with good results and enter a lucrative workforce (Talwar et al., 2019). The physiological and behavioral reactions associated with taking written or oral exams, which are felt both before and during the test due to concerns about potential unfavorable outcomes or exam failure, are referred to as test anxiety (Daniel & Ogunyewo, 2018; Jayasankara et al., 2018;).

To measure the test anxiety impairments in students, Driscoll (2004) created the Westside Test Anxiety Scale (WTAS), a concise 10-item scale designed to be completed in a short duration of five to eight minutes, providing an efficient measure of test anxiety. WTAS is a highly efficient test anxiety assessment instrument designed to find students who struggle with anxiety (Richard et al., 2023). It has been the primary anxiety measure in many studies and has been utilized for many years globally by school counsellors (Talwar, 2019).

The instrument has been used in many countries. In some African countries, such as Kenya, the WTAS has been adopted as a standard measure of test anxiety in many institutions because it is brief and effective (Richard et al., 2023). Countries like Nigeria, Ghana, Egypt, and South Africa have also used WTAS to measure test anxiety in educational settings (Samuel, 2021). In Nigeria, the WTAS has been widely applied in various educational contexts. However, most studies (Abiodun et al., 2017; Muhammad & Maimuna, 2018; Sani, 2021) have focused on measuring test anxiety and its related constructs rather than critically examining the scale's psychometric properties, especially at the item level. This leaves critical gaps in understanding its functionality and limitations. The current study addresses this gap by employing the Rasch model to provide an in-depth evaluation of the WTAS in the Nigerian context.

Moreover, the WTAS is unique compared to other psychometric tools due to its specific focus on cognitive and emotional aspects of test anxiety, its brevity, and its alignment with contemporary Rasch measurement principles. Unlike broader anxiety measures, the WTAS is designed specifically for assessing test anxiety in educational contexts, making it more targeted and practical for academic settings. Additionally, its unidimensional structure ensures that it provides a clear and focused measurement of test anxiety, distinguishing it from other multidimensional anxiety scales that assess general psychological distress rather than test-specific anxiety. Furthermore, the scale's ease of administration and scoring make it particularly useful for large-scale assessments and research applications.

The Rasch measurement model was chosen for this study due to its unique advantages over traditional psychometric methods in evaluating the psychometric properties of instruments. Unlike Classical Test Theory (CTT), which focuses on test-level reliability and item discrimination, the Rasch model provides detailed item-level analysis. It allows for an evaluation of specific items' performance in terms of difficulty, person-item fit, and test functioning on a common interval scale, making it particularly suitable for identifying issues such as misfitting items or gaps in the measurement range at the item level. Additionally, the Rasch model ensures that the scale operates invariantly across different subgroups, thereby addressing fairness, a critical component in test evaluation.

In this study, the Rasch model primarily assesses aspects of construct validity by evaluating Baghaei's (2008) framework on WTAS and the unidimensionality of the WTAS. While unidimensionality is a foundational requirement for construct validity, it alone may not be sufficient to fully establish this form of validity. Future research could extend this work by examining additional aspects of construct validity, such as convergent or discriminant validity, to provide a more comprehensive evaluation. The choice to focus on construct validity over content or criterion validity was driven by the study's aim to explore the underlying structure of the WTAS and its ability to measure the intended construct accurately.

Baghaei's (2008) framework was employed to analyze the evidence supporting the construct validity of the WTAS. Within this framework, two critical threats to construct validity —namely, construct-irrelevant variance and construct underrepresentation—were identified and investigated. Construct-irrelevant variance refers to the presence of undesired factors within the measurement that may distort the assessment of the intended construct. This issue was examined using infit and outfit mean square (MNSQ) statistics in the model-data fit analysis, which helps detect any misfit between the model and the data. On the other hand, construct under-representation pertains to the inability of the measure to capture essential aspects of the intended construct. To evaluate this threat, the study visually inspected the ordering of item difficulties. A smooth

and continuous progression in the item difficulties suggests that the scale adequately covers the construct without any significant gaps or areas of under-representation. These two concepts, examined within the context of Rasch model analyses, are crucial for ensuring the comprehensive and accurate measurement of test anxiety.

Psychometric evaluation encompasses more than reliability and validity; it also addresses fairness, various forms of validity, and diverse analytical approaches. This study acknowledges the importance of fairness by discussing how the WTAS performs across different subgroups, referencing Nnenanya et al.'s (2022) findings on invariance by gender and location in Kano State, Nigeria. While the current study emphasizes item-level psychometric properties through Rasch model analysis, it also highlights the need for further research into construct validity, particularly in understanding how well the WTAS measures test anxiety across diverse educational and cultural settings.

The integration of reliability and validity is strengthened by employing the Rasch model, which goes beyond traditional methods to offer detailed insights into item functioning and test structure. This modern psychometric approach complements traditional analyses, revealing critical gaps such as the scale's inability to adequately capture the abilities of high-performing students. By leveraging the strengths of modern methods like Rasch analysis, this study contributes to a nuanced understanding of the WTAS and provides a foundation for its refinement and contextual adaptation.

LITERATURE REVIEW

Due to the WTAS's popularity, numerous investigations into its psychometric properties have been carried out in a variety of contexts, with much of the research having produced positive results. However, many of these studies (Driscoll, 2004; Richard et al., 2023; Talwar et al., 2019) only addressed statistics at the test level, like factor structures and the internal consistency of the WTAS, neglecting a more detailed examination of the scale's psychometric properties at the item level. In addition, without item-level analysis, educators may miss opportunities to refine or eliminate poorly performing questions.

In Nigeria, several studies have applied WTAS to explore test anxiety across various educational settings, yet few have critically examined its psychometric properties. For instance, Nnenanya et al. (2022) conducted a study on the invariance of the WTAS by gender and location among secondary school students. This study revealed that the scale performed consistently across different demographic groups, suggesting that it is a robust tool for assessing test anxiety. However, Nnenanya et al.'s (2022) work primarily focused on the differential item functioning (DIF) between male and female students, without providing an in-depth evaluation of the scale's item-level psychometric properties, such as item difficulty or person-item fit. This leaves a significant gap in understanding the scale's capacity to measure test anxiety accurately at the item level within diverse student populations.

Sani (2021) further applied the WTAS to assess test anxiety among university students in Kano State, Nigeria. This study primarily explored the relationship between test anxiety and academic performance. While it confirmed that test anxiety was a significant factor influencing academic outcomes, the study did not address the underlying structure of the WTAS or its reliability and validity in measuring test anxiety among Nigerian university students. The absence of psychometric analysis in this study limits the generalizability of the findings, as it is unclear whether the WTAS accurately captures the construct of test anxiety in this specific context. In a similar vein, Abiodun et al. (2017) used the WTAS to evaluate the role of academic self-confidence in relation to test anxiety among Nigerian university students. Their study demonstrated that test anxiety negatively impacted academic self-confidence, but like the others, it did not explore the scale's psychometric robustness. The WTAS was used as a tool to assess anxiety. However, its item-level functioning and measurement properties were not considered, leaving room for improvement in understanding the scale's precision and reliability in the Nigerian context.

Additionally, Muhammad and Maimuna (2018) examined the impact of cognitive restructuring counseling approaches on test anxiety in secondary school students and other higher education institutions in Kano State, Nigeria. While this study is valuable for understanding intervention strategies, it similarly relied on the WTAS without addressing the scale's psychometric characteristics. The study did not explore whether the WTAS accurately measured test anxiety across different levels of educational attainment or whether any items were biased or misfitting. Despite the widespread use of the WTAS in Nigeria, there is a notable absence of research focused on evaluating its psychometric properties at the item level. This study aims to fill this gap by applying the Rasch model to assess the scale's construct validity, reliability, and fairness. The Rasch model allows for a more granular examination of the scale's item difficulty and person-item fit, providing a more nuanced understanding of its effectiveness in measuring test anxiety among Nigerian undergraduates. By addressing these psychometric concerns, this study offers critical insights into the strengths and weaknesses of the WTAS, ultimately contributing to its refinement for use in educational settings.

Driscoll's (2007) original development of the WTAS involved factor analysis to identify underlying dimensions of test anxiety and analysis at the test level to evaluate reliability and validity. While these methods are foundational in psychometric research, they primarily focus on the overall structure and consistency of the test, offering limited insights into individual item performance or the interaction between items and respondents. Traditional approaches, such as factor analysis, assume equal item functioning and do not account for the hierarchical relationship between item difficulty and respondent ability.

The WTAS offers a unique contribution to the assessment of test anxiety from a theoretical perspective. Unlike traditional multidimensional test anxiety scales, the WTAS is grounded in a unidimensional framework that aligns with cognitive interference theory, emphasizing the direct impact of anxiety on test performance. Its concise structure enables a focused measurement of test anxiety severity, distinguishing it from broader anxiety measures that assess multiple dimensions such as worry, emotionality, and physiological responses. Additionally, this study provides new empirical evidence on the psychometric properties of the WTAS using Rasch model analysis, further validating its theoretical foundation. By demonstrating its effectiveness in measuring test anxiety within a Nigerian undergraduate context, this study highlights the WTAS's relevance as a robust and efficient tool for assessing test anxiety in educational settings.

This study is grounded in Item Response Theory (IRT), particularly the Rasch model, which provides a robust framework for evaluating the psychometric properties of the WTAS. The Rasch model ensures that measurement scales adhere to fundamental principles of construct validity, item functioning, and measurement invariance, offering a more refined analysis of test anxiety assessments. By applying Rasch analysis, this study contributes to the theoretical discourse on how psychological constructs such as test anxiety can be measured with precision and reliability.

The Rasch measurement model offers significant advantages over these traditional methods. Unlike factor analysis, the Rasch model operates at the item level, providing detailed information on item difficulty, discrimination, and fit. It transforms ordinal raw scores into interval-level measurements, enabling more precise interpretations of test anxiety levels. Additionally, the Rasch model evaluates the invariance of item functioning across subgroups, ensuring that the test measures the construction equitably. This is particularly important in diverse educational settings like those in Nigeria, where differences in demographic and cultural factors may influence test performance.

Furthermore, the Rasch model offers a Wright map that visually aligns respondent abilities with item difficulties, highlighting gaps in the measurement range and identifying items that may not align with any respondent abilities. This level of detail provides actionable insights for refining the WTAS, such as adding items to better capture high or low levels of test anxiety. By employing the Rasch model, this study validates the WTAS and contributes to its enhancement, addressing limitations that traditional methods cannot resolve.

On the other hand, Classical Test Theory (CTT) posits that an individual's score comprises both their true ability score and some measurement error. It is important to note that neither the true score nor the error score can be directly observed. Consequently, certain assumptions are necessary to evaluate students' scores. These

assumptions were outlined by Hambleton and Jones (1993) as follows: (a) the absence of correlation between unobserved scores, (b) a zero mean error score in the population, and (c) the absence of correlation between error scores from parallel tests. The current study seeks to assess the psychometric qualities of the WTAS by employing the Rasch measurement model analysis. This assessment is essential because understanding the psychometric qualities of the WTAS will enable researchers to obtain accurate, reliable, and valuable data, thereby facilitating a more appropriate interpretation of the results. This, in turn, improves research quality, provides information for focused interventions and policies, and contributes to the body of knowledge in education.

The advent of the modern test theory (IRT), along with associated measurement models, has enabled a more precise analysis of dichotomously scored data. Specifically, the Rasch Measurement Model, a subset of IRT, provides a framework for estimating an individual's score using Equation 1, as described by Khairani et al. (2020):

$$P(\theta) = \frac{\exp(\theta - b)}{1 + \exp(\theta - b)} \quad [1]$$

Where: P(θ) = is the probability of a correct response (x=1) given the person's ability (θ); b = represents the item difficulty or threshold; θ = represents the person's ability or trait level; Exp = is the exponential function

Furthermore, the Rasch Model is a renowned measurement framework that facilitates the conversion of ordinal raw scores into equivalent interval measures, thereby enabling linear measurement (Linacre, 2006). This transformation allows for the estimation of precise and comparable measures, fostering a more accurate understanding of the underlying construct. As a result, this approach is increasingly being used to validate instruments (Dina et al., 2024; Emerson et al., 2022; Farshad & Purya, 2024; Hamad, 2021; Khairani et al., 2020; Mohd Matore & Khairani, 2020; Mustapha & Ehab, 2022; Nadhirah et al., 2022; Richard et al., 2023). Rasch model analysis, in contrast to CTT, is more resilient to the distribution's normality assumption and may even determine a person's score if some data is absent. Furthermore, as opposed to test-level statistics like those in the CTT, the Rasch model analysis offers a more nuanced understanding of the data by providing item-level statistics, which facilitate a deeper interpretation of the results. These statistics enable researchers to examine the performance of each item, identifying areas where the measure may be improved, and ultimately leading to a more comprehensive understanding of the construct being measured. This includes more comprehensive proof of the instrument's psychometric properties, like validity and reliability.

In addition, IRT offers insights into the testing's fairness, i.e., whether subgroups of the sample being tested (such as location, gender, and socioeconomic status) perceive the instrument's items similarly. Moreover, it is stated that the framework can assess whether the tool being used can effectively

target the specified sample. Muis et al.'s (2009) study, which combined Classical Test Theory (CTT) and Rasch model analyses, revealed that while CTT analyses showed promising results regarding the instrument's reliability and validity, the Rasch model analysis uncovered significant limitations and issues that were not apparent through CTT. Based on these findings, we concluded that the Rasch model is a more comprehensive and robust analytical approach compared to CTT, providing a more accurate understanding of the instrument's measurement properties.

Likewise, Baghaei's (2008) framework offers a robust approach to evaluating the construct validity of measurement instruments by addressing two critical threats: construct-irrelevant variance and construct underrepresentation. Construct-irrelevant variance refers to the inclusion of factors unrelated to the intended construct, such as respondent characteristics or external influences, which may distort the measurement results. On the other hand, construct underrepresentation occurs when the measurement instrument fails to capture all critical aspects of the intended construct, leading to an incomplete assessment. By addressing these two threats, Baghaei's (2008) framework ensures that instruments provide a valid and comprehensive measurement of the construct of interest, aligning with the principles of psychometric validity.

In the context of Rasch model analysis, Baghaei's (2008) framework is particularly useful for evaluating model-data fit and identifying potential issues with the measurement instrument. For instance, construct-irrelevant variance can be detected through infit and outfit mean square statistics, which highlight items that do not conform to the expected measurement model. Similarly, construct underrepresentation can be assessed using the item-person map, which examines the alignment between item difficulties and respondent abilities to ensure comprehensive coverage of the construct. This framework provides a theoretical foundation for the current study, guiding the analysis of the WTAS and ensuring that the scale accurately and comprehensively measures test anxiety. Including this perspective enhances the validity and reliability of the study findings.

METHODOLOGY

Research Design

Data for the present study were gathered over a single time using a cross-sectional study method. This approach was selected due to its capacity to collect a large amount of data about the psychometric properties of the WTAS from an extensive sample of respondents.

Respondents

A total of 300 undergraduates from the faculties of Education, Humanities, Natural and Applied Sciences, Management Sciences, Agriculture, and Information and Communication Technology (ICT) at Sule Lamido University participated in this study. The convenience sample consisted

of 300 participants, including 160 females (53.3%) and 140 males (46.7%), who were selected during lecture classes based on their accessibility and willingness to participate. The average age of the respondents ranged from 19 to 30 years old, and they were all in their first year at university.

Measures

The study utilized the 10-item WTAS questionnaire, which is formatted on a Likert scale designed to detect and evaluate test anxiety impairment in students. Example items from the WTAS include statements such as: *"The closer I am to a major exam, the harder it is for me to concentrate on the material," "When I study, I worry that I will not remember the material on the exam,"* and *"During important exams, I think that I am doing awful or that I may fail."* According to Driscoll (2007), WTAS can be scored as follows: Scores between 1.0 and 1.9 indicate low test anxiety, 2.0 and 2.5 represent average test anxiety, 2.5 and 2.9 indicate higher than average test anxiety, 3.0 and 3.4 suggest moderately high test anxiety, with some items indicating high levels of anxiety, 3.5 and 3.9 indicate high test anxiety, with half or more items rated as high, and 4.0 and 5.0 reflect extremely high test anxiety, with items rated as high or extreme.

Data Collection and Analysis

The Sule Lamido University administration granted written consent for the purpose of data collection. Prior to administering the scale, students were fully informed about the study's objectives, thereby guaranteeing that their responses would be treated with utmost confidentiality and anonymity, ensuring a secure and private assessment experience. The questionnaire in the form of a Google form was shared with the students' WhatsApp groups during the last 30 minutes of the 2-hour lecture classes. Participation was voluntary and applicable only to students who were present during the assessment. The study received university approval.

Rasch model analysis was employed to analyze the collected data in the following order: (1) Rasch model's assumptions, (2) evidence of construct validity and reliability, (3) differential item functioning (DIF) analysis between male and female students, (4) the item person map (Wright map). Strict assumptions are essential when using the Rasch Model as a measurement framework. The two important Rasch model assumptions are as follows: (1) The data must demonstrate a good fit with the model's expectations, indicating a compatible and coherent relationship and (2) The construct being measured must possess a unidimensional nature, meaning it can be represented by a single underlying trait or dimension, without any additional underlying factors influencing the measurements, Linacre (2006).

The assumption of the Rasch model's data fit was evaluated using the infit and outfit mean squares (MNSQ) statistics, which provide a measure of how well the data aligns with the Rasch model's expectations. According to Bond and Fox

(2015), MNSQ values between 0.6 and 1.4 logits indicate an acceptable fit, suggesting that the data is behaving in line with the model's assumptions. Additionally, the assumption of unidimensionality was assessed through Principal Component Analysis (PCA) of residual procedures, which helps to identify any potential additional dimensions or factors that may be present in the data. By examining the results of these analyses, we can determine whether the data meet the necessary assumptions for Rasch model analysis, ensuring the validity and accuracy of our measurements. This technique aims to determine whether other components can be defined from the residuals by removing the most significant factor. An eigenvalue greater than 2.0 for the first extracted construct is considered to violate the unidimensional assumption (Linacre, 2006). According to the same author, a scale demonstrates unidimensionality when the primary measurement dimension accounts for greater than 20% of the total variance, indicating that the scale effectively measures a single underlying construct.

The Rasch measurement analysis offers a clear statistical method for evaluating the measurement's reliability. Information about the reproducibility of the results can be found in the item difficulty reliability index. This index is calculated by dividing the observed variance in item difficulty by the true variance in item difficulty. Items with higher reliability indices are more likely to accurately reflect their true difficulty levels, ensuring that items identified as more difficult genuinely possess higher difficulty compared to those with lower difficulty measurements. The sample size and difficulty variance determine the item's difficulty reliability. High item difficulty reliability values are the result of large sample sizes and a wide range of item difficulties, and vice versa. Values greater than 0.80 are regarded as acceptable by Bond and Fox (2015); however, Fisher (2007) classifies values greater than 0.94 as strong.

Baghaei's (2008) framework was used to analyze the evidence supporting the construct validity of this measure. Construct-irrelevant variance and construct underrepresentation are the two threats to construct validity that need to be identified to investigate construct validity, according to this approach. The question of whether the measurement contains undesired constructions is at the heart of construct-irrelevant variance. It was found by utilizing the previously described infit and outfit MNSQ in the model-data fit analysis. However, the term "construct underrepresentation" describes the measurement's inability to capture significant facets of the desired concept. The threat of construct under-representation was evaluated by visually examining the ordering of the items' difficulties, where a continuous and smooth progression between items indicates that the scale is comprehensively capturing the construct, without any notable gaps or weaknesses.

Summary statistics were calculated to assess the reliability and separation indices for both respondents and items. These analyses aimed to evaluate the scale's ability

to differentiate between respondent ability levels and to measure the precision and consistency of item difficulty estimates. This approach ensures a thorough examination of the WTAS's psychometric properties and its capacity to reliably measure test anxiety across diverse samples.

In the Rasch Model's item-level analysis, another important statistic is the Differential Item Functioning (DIF) analysis. DIF analysis examines whether a particular item favors one group over another. If different groups perceive the same item differently, it suggests variation in interpretation. According to Bond and Fox (2015), a DIF contrast statistic greater than 0.5 logits indicate evidence of DIF items. In the present study, we investigated whether there were differences in how boys' and girls' respondents perceived the items, as reported by numerous past studies (Hamad, 2021; Mohd Matore & Khairani, 2020; Mustapha & Ehab, 2022; Richard et al., 2023).

RESULTS

The results showed that the mean infit MNSQ was 1.00 logits (SD = 0.07 logits), and the mean outfit MNSQ was 0.99 logits (SD = 0.08 logits), revealing that the items generally fit the expected value of 1.00 logits (Table 1). Furthermore, the infit MNSQ values for each item ranged from 0.90 to 1.10 logits, and the outfit MNSQ values ranged from 0.85 to 1.08 logits, all falling within the acceptable range of 0.6 to 1.4 logits recommended by Bond and Fox (2015). These results suggest that the items demonstrate a good fit to the Rasch model, with no evidence of misfit or unexpected behavior.

As presented in Table 2, the eigenvalue for the primary construct was 1.8, which falls below the recommended threshold of 2.0 suggested by Linacre (2006) for establishing unidimensionality. However, it's important to note that this primary construct accounts for 28.5% of the variance in the measurement, indicating that the majority of the variance is explained by the first dimension, supporting the assumption of unidimensionality.

The summary statistics presented in Tables 3 and 4 confirm the strong psychometric properties of the WTAS. Table 3 shows a person reliability of 0.73, indicating acceptable consistency in respondents' performances. Additionally, the separation index of almost 2 reflects the scale's ability to distinguish respondents

Table 1
Descriptive statistics (in logits)

Variables	Measure	Model Error	Infit MNSQ	Outfit MNSQ
Mean	0.00	0.07	1.00	0.99
SD	0.33	0.00	0.07	0.08
Max.	0.40	0.07	1.10	1.08
Min.	-0.38	0.06	0.90	0.85

into at least two distinct levels of test anxiety. In addition, Table 4 reports a high item reliability of .96 and a separation index of 4.92, demonstrating the scale's ability to differentiate item difficulties with great precision. These findings validate the WTAS's measurement quality and highlight its reliability in assessing test anxiety.

Table 2
Findings from the Principal Component Analysis (PCA) of residuals

		Empirical		Modelled	
Total raw variance observed	=	40.0	100.0%		100.0%
Raw variance explained by measures	=	4.0	28.5%		28.8%
Raw variance explained by individuals	=	1.3	9.2%		9.3%
Raw Variance explained by the items	=	2.7	19.3%		19.5%
Raw unexplained variance (total)	=	10.0	71.5%	100.0%	71.2%
Unexplained variance in the first contrast	=	1.8	12.5%	17.5%	
Unexplained variance in the second contrast	=	1.3	9.6%	13.5%	
Unexplained variance in the third contrast	=	1.2	8.5%	11.9%	
Unexplained variance in the fourth contrast	=	1.1	8.1%	11.3%	
Unexplained variance in the fifth contrast	=	1.1	7.9%	11.0%	

Table 3
Summary statistics: Person

	Total Score	Count	Measure	Model Error	Infit		Outfit	
					MNSQ	ZSTD	MNSQ	ZSTD
Mean	27.8	10.0	0.33	0.38	0.99	0.0	0.99	0.0
SD	5.3	0.0	0.78	0.16	0.42	1.2	0.45	1.2
Max.	38.0	10.0	2.19	1.84	2.27	2.4	2.74	2.6
Min.	10.0	10.0	-4.25	0.33	0.09	-4.4	0.09	-4.2
Real RMSE		0.43	True SD	0.65	Separation	1.51	Person Reliability	0.69
Model RMSE		0.41	True SD	0.67	Separation	1.65	Person Reliability	0.73

Table 4
Summary statistics: Item

	Total Score	Count	Measure	Model Error	Infit		Outfit	
					MNSQ	ZSTD	MNSQ	ZSTD
Mean	834.3	300.0	0.00	0.07	1.00	0.1	0.99	-0.1
SD	77.9	0.0	0.33	0.06	0.07	1.0	0.08	1.0
Max.	965.0	300.0	0.40	0.07	1.10	1.4	1.08	1.0
Min.	732.0	300.0	-0.62	0.06	0.90	-1.3	0.85	-1.06
Real RMSE		0.07	True SD	0.32	Separation	4.83	Item Reliability	0.96
Model RMSE		0.07	True SD	0.32	Separation	4.92	Item Reliability	0.96

Furthermore, the item reliability measure of 0.96 is considered an acceptably strong value, as noted by Bond and Fox (2015). The results also demonstrate strong construct validity, as all the items showed a good fit between the empirical data and the Rasch model's expectations, with acceptable infit and outfit MNSQ values. This indicates that the items are measuring the intended construct without significant interference from irrelevant variables, thereby minimizing the risk of construct-irrelevant variance. Inspections of item difficulty measurements between items are presented in Table 5. According to Linacre (2006), the gaps between consecutive items are all less than 0.5 logits, indicating that the gaps are relatively small and likely insignificant. This suggests that the items are evenly spaced and that the measurement scale is continuous, without any notable gaps or discontinuities. The table further presented that, according to the acceptable DIF contrast values of 0.00 - 0.23 logits, both genders interpreted the items similarly and the items do not favor any group.

On the other hand, the Wright Map visually represents the distribution of persons and items on the measurement ruler in logits, offering a comprehensive view of both the abilities of individuals and the difficulties of items. As shown in Figure 1, the respondents are represented by #, and items are labeled Item1–Item10. "." indicates the proportion of the three respondents, and each # represents three respondents. The result shows that the most difficult item is Item 4 (+0.40 logits), and the easiest item is Item 5 (-0.62 logits). The logits ranged from +0.40 to -0.60 logits, and that satisfied the acceptable range of +3.0 to -3.0 logits that was considered acceptable (De Klerk et al., 2013; Linacre, 1994).

Table 5
Item statistics and differential item functioning (DIF) contrast (in logits)

Item	Measure	Count	SE	Infit MNSQ	Outfit MNSQ	DIF Contrast
4	732	300	0.06	1.09	1.07	0.04
10	751	300	0.06	0.96	.93	-0.30
7	790	300	0.06	1.03	1.02	0.05
9	793	300	0.07	0.91	.89	-0.06
1	824	300	0.06	1.05	1.06	-0.26
6	801	300	0.07	0.97	.97	0.20
3	814	300	0.07	1.08	1.05	-0.07
2	943	300	0.07	0.90	.85	0.23
8	930	300	0.07	1.10	1.08	0.09
5	965	300	0.07	0.94	.97	0.00
Mean	0.00	300.0	0.07	1.00	.99	
SD	0.33	0.0	0.00	0.07	.08	

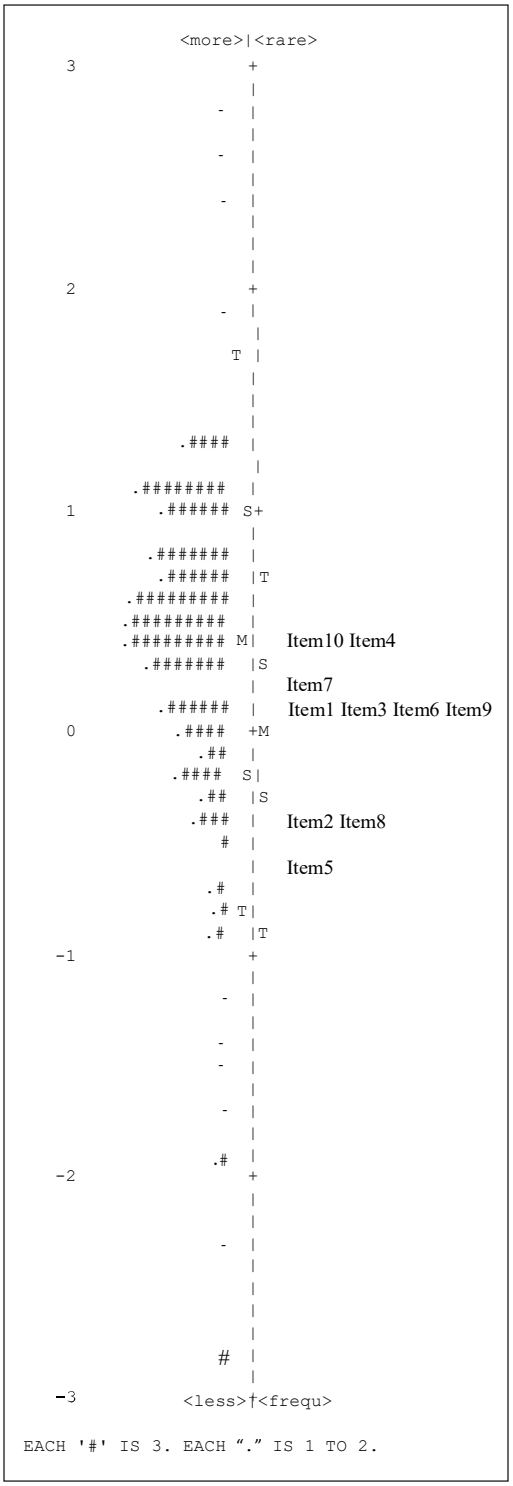

Figure 1. Item Person Map (Wright Map)

DISCUSSION

The primary objective of this study is to investigate the psychometric qualities of the WTAS using a sample of Nigerian undergraduates, which is one of the successful instruments created to evaluate test-takers' anxiety levels. To achieve this goal, the study assessed the WTAS using the criteria within the Rasch Model analysis framework, including the Rasch model's assumptions, reliability, evidence of construct validity, Differential Item Functioning (DIF), and the Wright Map. The results from the fit statistics and Principal Component Analysis (PCA) of residuals showed that all assumptions of the Rasch model were met, revealing a good fit between the data and the model. Specifically, the fit statistics demonstrated that the items were functioning as expected, and the PCA results confirmed the unidimensionality of the scale, with no significant residual variance. Therefore, the measurement of test anxiety using the WTAS demonstrates properties of equal-interval measurement. This finding was also supported by Richard et al. (2023), who analyzed the psychometric qualities of the WTAS with a sample of students from Kenya (Africa). They employed the Rasch model criteria, and all the assumptions were also met. Since the findings of the present study indicate that the Westside test anxiety scale is a valid and reliable measure of test-anxiety impairment in a sample of Nigerian undergraduates, we therefore recommend WTAS for use in Nigeria.

Furthermore, the encouraging outcome from the good reliability value in the present

study may be attributed to the large coverage of item difficulty and the sample size that was employed. Therefore, for a similar sample of test takers, it may be concluded that the difficulty of WTAS items is highly replicable. This means that there is a good chance the item ordering in the WTAS won't change if it is given to a different sample of Nigerian university students. The study conducted in Malaysia by Talwar et al. (2019) likewise reported on the good reliability of the WTAS items.

Meanwhile, the WTAS reports positive results for construct validity evidence, indicating little evidence of construct-irrelevant variation and construct under-representativeness. Given that the WTAS measures test anxiety impairment and no other unintended constructs, test users can interpret test results with confidence according to the excellent evidence of the construct validity of the instrument. The outcome is not unexpected at this point, as other research conducted across a wide range of cultural contexts has demonstrated the WTAS's strong construct validity (Ehab, 2022; Richard, 2023; Talwar, 2019). Moreover, Driscoll's (2004) initial validation of the WTAS provided evidence of strong construct validity and reliability at the test level, making it a widely accepted tool for assessing test anxiety. However, the original analysis primarily focused on overall test performance without delving into item-level psychometric properties. Specifically, key aspects such as item difficulty, fit, and functionality across varying respondent abilities were not addressed. Furthermore, the WTAS has been widely used in different cultural and educational settings, yet there is limited evidence of its performance in contexts outside its original development environment, particularly in Nigerian universities. This study addresses these gaps by employing the Rasch model to evaluate the WTAS at the item level, providing insights into its construct validity in a new context. The DIF analysis shows that the items do not differ between genders, and this finding is supported by various studies (Muhammad & Maimuna, 2018; Nnenanya et al., 2022).

The Wright map analysis revealed that the abilities of approximately half of the students exceeded the difficulty levels of the items on the WTAS. This finding indicates that many students possess higher test-taking abilities compared to the demands of the items in the scale. Such an alignment suggests that while the WTAS effectively captures a range of test anxiety levels, it may lack sufficient items targeting respondents with higher abilities. This mismatch highlights a potential ceiling effect, where the existing items are less informative for students with abilities significantly above the average difficulty level. Addressing this gap could involve revising or expanding the scale to include items that better differentiate among high-ability respondents, thereby enhancing the WTAS's ability to provide more comprehensive measurement across the entire spectrum of test-taker abilities.

Additionally, the WTAS is a concise ten-item scale specifically developed to evaluate and identify students experiencing test anxiety

impairment, this study provides evidence that the high performing respondents were not effectively targeted by the WTAS, particularly 142 respondents at the upper and 15 respondents at the lower ends of the scale continuum (Figure 1) the respondents are far from the items. Therefore, we recommend exploring alternative strategies to ensure the scale adequately captures the abilities of respondents across the entire spectrum, particularly those at the extreme ends of test anxiety.

The findings of this study contribute to the body of knowledge by providing empirical evidence on the psychometric properties of the WTAS, as assessed using the Rasch model. This approach has been largely unexplored at the item level. The study highlights the scale's limitations in differentiating high-ability respondents, emphasizing the need for item refinement to enhance measurement precision. Additionally, by examining WTAS within a Nigerian undergraduate population, this research extends its validity and applicability to diverse educational contexts, offering valuable insights for future test development and adaptation.

CONCLUSION

Using the Rasch measurement model analysis, this study has demonstrated the strong psychometric properties of the 10-item WTAS among Nigerian university students. This study has supplied empirical data on the items. The results indicated a need to revise or expand the item pool to better assess anxiety in students with high abilities, ensuring that the scale is comprehensive and inclusive for all students, regardless of their ability level. Future studies should proactively test the items with students from other postsecondary educational institutions, like colleges of education and polytechnics. Only university students are included in this research. Research should go beyond psychometric issues to have a deeper understanding of the WTAS items. Therefore, researchers need to conduct more qualitative research to determine what factors contribute to one item's difficulty relative to others, which raises the level of test anxiety.

Implications of the Study for Theory and Practice

The results of our Rasch model analysis of the WTAS have significant implications for both theory and practice. Firstly, this study contributes to the theoretical understanding of test anxiety measurement, particularly within the Nigerian context. By employing the Rasch model to evaluate the WTAS's psychometric properties, we establish its construct validity, reliability, and differential item functioning (DIF), reinforcing its suitability for assessing test anxiety. The findings also underscore the Rasch model's robustness in validating psychological constructs, demonstrating that it provides a more precise, item-level evaluation compared to classical test theory (CTT). This highlights the necessity of employing modern measurement techniques to ensure that assessment instruments accurately reflect the latent traits they intend to

measure. Additionally, our study reveals the importance of cultural adaptation in psychometric testing, emphasizing that test anxiety manifestations may vary across populations. This insight encourages further refinement of psychological assessment tools to account for cultural and contextual differences, ultimately improving their generalizability and applicability across diverse student populations.

Secondly, from a practical perspective, this study provides valuable insights for educators, psychologists, and counselors who seek to identify and support students struggling with test anxiety. The strong reliability and validity of the WTAS suggest it can be effectively used in educational and clinical settings to screen for test anxiety and develop targeted interventions. However, the Wright map analysis indicates that the abilities of approximately half of the students are above the difficulty level of the items, suggesting that the scale may not fully differentiate among students with lower anxiety levels. To enhance its effectiveness, practitioners should consider complementing the WTAS with additional diagnostic measures, such as qualitative assessments, structured interviews, or physiological indicators of anxiety. Furthermore, the study underscores the need for interventions that specifically address varying levels of test anxiety, particularly among high-performing students who may experience anxiety in subtler yet academically impactful ways. Schools and universities can use these findings to implement structured support programs, including cognitive-behavioral strategies, relaxation techniques, and test-taking skills training. Ultimately, integrating these evidence-based approaches into educational policies and counseling services can help mitigate test anxiety's negative effects, fostering improved academic performance and student well-being.

Limitation of the Study

Despite its valuable findings, this study has several limitations that warrant acknowledgment. The use of convenience sampling from a single public university in Nigeria limits the generalizability of results to the broader student population across various institutions and regions. Furthermore, the focus on first-year students excludes understanding of test anxiety patterns among students in later years of study. The cross-sectional design prevents causal inferences, necessitating longitudinal studies to explore the stability of WTAS scores over time and understand how test anxiety evolves throughout students' academic journeys.

While the study provides insights into the WTAS's psychometric properties in Nigeria, cultural differences may affect its applicability elsewhere, requiring further research in diverse settings. The Wright map analysis revealed that WTAS items inadequately target respondents at the extreme ends of the test anxiety spectrum, suggesting the scale may not fully capture the range of experiences, particularly for students with very high or low anxiety levels. The reliance on self-reported data

introduces potential response biases, which could be mitigated in future studies by incorporating objective measures or behavioral observations. Additionally, the lack of comparisons with other established test anxiety measures limits the evidence for the WTAS's concurrent validity. Lastly, the study did not extensively examine the impact of demographic variables on test anxiety, leaving room for a more nuanced understanding of how different factors influence anxiety levels among students. Addressing these limitations in future research will enhance the robustness of the WTAS and its applicability in diverse educational settings.

Recommendations for Future Studies

Based on the findings and limitations of this study, we propose several directions for future research. Future studies should include a more diverse and representative sample of students from various Nigerian universities and regions to enhance the generalizability of findings and provide a more comprehensive understanding of test anxiety across different contexts. Conducting longitudinal research would allow examination of WTAS score stability over time, offering insights into how test anxiety evolves throughout students' academic careers and potentially identifying the long-term effectiveness of anxiety-reduction interventions. To ensure the WTAS's applicability in diverse cultural contexts, further research should validate the scale in different countries and educational environments, making cross-cultural comparisons to determine its universal applicability and identify any necessary culture-specific adaptations.

Given that the Wright map analysis revealed limitations in targeting respondents at the extreme ends of the anxiety spectrum, future studies should develop revised or expanded items to better capture the experiences of students with very high-test anxiety. To address potential biases in self-reported data, future research should incorporate objective measures or behavioral observations, combining self-report scales with physiological measures or performance-based assessments. Including comparisons with other established test anxiety measures will provide additional evidence for the WTAS's concurrent validity. Furthermore, future studies should explore the impact of various demographic variables on test anxiety and focus on developing and testing interventions designed to reduce test anxiety. By addressing these recommendations, future research can build on the current study's findings, enhance the robustness of the WTAS, and contribute to a more comprehensive understanding of test anxiety in diverse educational settings.

ACKNOWLEDGEMENT

The authors are deeply grateful for the support and assistance received from the School of Educational Studies, Universiti Sains Malaysia, Malaysia, and the School of General and Entrepreneurship Studies, Sule Lamido University Kafin Hausa, Nigeria, which enabled us to successfully complete this research paper. Their contributions are sincerely appreciated.

REFERENCES

Abiodun, M. L., Erhabor, S., & Omotola, P. A. (2017). Academic self-confidence effects on test anxiety among Nigerian university students. *Journal of Psychology in Africa, 28*(7), 307-315.

Baghaei, P. (2008). The Rasch Model as a construct validation tool. *Rasch Measurement Transactions, 22*(1), 1145–1146.

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: The fundamental measurement in human science* (3rd ed.). Routledge. https://doi.org/10.4324/9781315814698.

Emerson, B., Iqbal, N., & Medvedev, O. (2022). Enhancing the assessment of gratitude in mindfulness research: A Rasch analysis of the 6-item Gratitude Questionnaire. *Mindfulness, 13*(12)*,* 3017-3027. https://doi.org/10.1007/s12671-022-02006-2

Daniel, C.O., Ajayi, A.D., Ogunyewo, O.A., Ajio, D.K., Andy, E., Oyedele, E., Daniel, G., Solomon, G.M., Haruna, A.B., & Gusen, N.J. (2018). Perceived impact of academic stress on academic performance among undergraduate nursing students of university of Jos. *New Nigerian Journal of Clinical Research*, *7*, 8-15.

Darman, D. R., Suhandi, A., Kaniawati, I., Samsudin, A., & Wibowo, F. C. (2024). Developmental validation of scientific inquiry literacy instrument (SILI) using Rasch measurement model. *Education Sciences, 14*(3), Article 322. https://doi.org/10.3390/educsci14030322

De Klerk, M., Nel, J.A., Hill, C., & Koekemoer, E. (2013). The development of the MACE work-family enrichment instrument. *SA Journal of Industrial Psychology, 39*(2), Article 1147. https://doi.org/10.4102/sajip.v39i2.1147

Driscoll, R. (2007). *Westside Test Anxiety Scale validation* (ED495968). American Test Anxiety Association.

Driscoll, R. (2004). Westside test anxiety scale. American test anxiety association. *Online Submission*. Retrieved from https://files.eric.ed.gov/fulltext/ED495968.pdf

Farshad, E., & Purya, B. (2024). Examining the dimensionality of linguistic features in L2 writing using Rasch measurement framework model. *Educational Methods & Practice, 2*. 1-22.

Fisher, W. P. (2007). *Rating scale instrument quality criteria*. Rasch Measurement Transactions*, 21*(1), 1095. https://www.rasch.org/rmt/rmt211m.htm

Hamad, A. A. (2021*). The psychometric properties of measurement of the mathematics teachers' professional identity in Saudi Arabia* [Doctoral dissertation, Universiti Sains Malaysia]. http://eprints.usm.my/52981/1/ALGHAMDI%20HAMAD%20ABDULLAH%20J%20-%20TESIS.pdf%20cut.pdf

Hambleton, R. K., & Jones, R. W. (1993). A comparison of the classical test theory and item response theory and their applications to test development. *Educational Measurement: Issues and Practice, 12*(3), 38-47. https://doi.org/10.1111/j.1745-3992.1993.tb00543.x

Jayasankara, R. K., Rajan, M. K., & Thattil, A. (2018). Academic stress and its sources among university students. *Biomedicine & Pharmacotherapy Journal, 11*(1), 531-537. https://doi.org/10.13005/bpj/1404

Keith, N., Hodapp, V., Schermelleh-Engel, K., & Moosbrugger, H. (2019). Cross-sectional and longitudinal confirmatory factor models for the German Test Anxiety Inventory. *Anxiety, Stress & Coping, 16*(3), 251-270. https://doi.org/10.1080/1061580031000095416

Khairani, A. Z., Matore, M., Effendi, M., & Khairani, M. Z. (2020). Psychometric properties of Raven's Advanced Progressive Matrices with a sample of Malaysian youth. *Pertanika Journal of Social Sciences & Humanities, 28*(1), 267-277.

Khalaf, M. A., & Omar, E, M. N. (2022). Rasch analysis and differential item functioning of English language anxiety scale (ELAS) across sex in Egyptian context. *BMC Psychology, 10*(1), Article 242. https://doi.org/10.1186/s40359-022-00955-w

Khan, E. A., Munir, H., Afzal, A., & Ansari, M. M. (2021). Influence of test anxiety on students' academic achievement at the secondary level. *Ilkogretim Online, 20*(2), 303-316. https://doi.org/10.17051/ilkonline.2020.02.31

Linacre, J. M. (1994). Sample size and item calibration stability. *Rasch Mes Trans*, *7*, 328.

Linacre, J. M. (2006). *A user's guide to WINSTEPS/ MINISTEPS: Rasch model computer programs*. Winsteps. https://www.winsteps.com/winman/copyright.htm

Messick, S. (1993). Validity. In R. L. Linn (Ed.), *Educational measurement* (2nd ed., pp. 13–104). American Council on Education and Oryx Press. https://www.scirp.org/reference/ReferencesPapers?ReferenceID=590541

Mohd Matore, M. E, E. & Khairani, A. Z. (2020). The Psychometric properties of Adversity Quotients (AQ) item for TVET education using Rasch Models. *Journal of Technical Education and Training*, 12(3), 61-70.

Muhammed, H. J., Mandaras, T., Aman, M., Tesfaye, T., & Wubiset, G. (2024). Test anxiety and its determinant among health science students at Mattu University: A cross-sectional study. *Frontiers of Psychiatry, 14.* https://doi.org/10.3389/fpsyt.2023.1241940

Muhammad, H. R., & Maimuna, M. S. (2018). An effect of cognitive restructuring counselling technique on test anxiety among students of tertiary institution' staff secondary students in Kano State of Nigeria. *Prestige Journal of Counselling Psychology, 1*(1), 2651-5709.

Muis, K. R., Winnie, P. H., & Edwards, O. V. (2009). Modern psychometric for assessing achievement of goal orientation: A Rasch analysis. *British Journal of Educational Psychology, 79*, 547-576. https://doi.org/10.1348/000709908x383472

Mustafa, A. K., & Ehab, M. N. (2022). Rasch analysis and differential item functioning of English language anxiety scale (ELAS) across sex in Egyptian context. *BMC Psychology, 10*(242), 4. https://doi.org/10.1186/s40359-022-00955-w

Nadhirah, N. A., Suryana. D., & Ahmad, A. (2022). Development and validation of peaceful classroom scale: Rasch model analysis. *International Journal of Instruction, 15*(4), 497-514.

Nnenanya, G. C., Okeke, A. O., Ezugwu, J. I., Nwani, S. U., & Eze, C. O. (2022). Invariance of the Westside test anxiety scale by sex and location among secondary school students. *African Academic Research Forum* (p. 193).

Panayides, P., Robinson, C., & Tymms, P. (2010). The assessment revolution that has passed England by: Rasch measurement. *British Educational Research Journal*, *36*(4), 611–626. https://doi.org/10.1080/01411920903018182

Pekrun, R., & Stephens, E. J. (2012). Academic emotions. In K. R. Harris, S. Graham, T. Urdan, S. Graham, J. M. Royer, & M. Zeidner (Eds.), *APA educational psychology handbook* (Vol. 2, pp. 3–31). American Psychological Association. https://doi.org/10.4236/eng.2011.34046

Putwain, D. W., & Daly, A. L. (2013). Does clusters of test anxiety and academic buoyancy differentially predict academic performance? *Learning and Individual Differences, 27*, 157–162. https://doi.org/10.1016/j.lindif.2013.07.010

Raufelder, D., & Ringeisen, T. (2016). Self-perceived competence and test anxiety: The role of academic self-concept and self-efficacy. *Journal of Individual Differences, 37*(3), 159–167. https://doi.org/10.1027/1614-0001/a000202

Richard, M. W., Peter, M. A., & Jotham, N. D. (2023). Psychometric properties of a test anxiety scale

for use in computer-based testing in Kenya. *The International Journal of Assessment and Evaluation, 30*(1), 2327-8692. https://doi.org/10.18848/2327-7920/CGP/v31i01/1-18.

Samuel, T. A. (2021). Assessing the validity of test anxiety tools in enhancing students' performance. *Journal of Educational and Psychological Measurement, 76(5*), 723-733.

Sanchez Rosas, J., & Furlan, L. (2017). The achievement emotions and achievement goals in support of the convergent, divergent and criterion validity of the Spanish-Cognitive Test Anxiety Scale. *International Journal of Educational Psychology, 6*(1), 67–92. https://doi.org/10.17583/ijep.2017.2268

Sani, S. (2021). Test anxiety among university students in Kano State Nigeria. *A Bi-lingual Peer-Reviewed Academic Journal, 1*(1), 2582-0427. https://doi.org/10.37948/ensemble-2021-CSI1-a019

Schnell, K., Ringeisen, T., Raufelder, D., & Rohrmann, S. (2020). The impact of adolescents' self-efficacy and self-regulated goal attainment processes on school performance – Do gender and test anxiety matter? *Learning and Individual Differences, 38*, 90–98. https://doi.org/10.1016/j.lindif.2014.12.008

Scopeliti, M., Tosi, J., & Zanetti, M. A. (2021). Assessing the unidimensionality of psychological scales: A comparison between Rasch model and factor analysis. *Psychological Test and Assessment Modeling*, *63*(1), 3–24.

Talwar, P., Matheiken, S., Cheng, J. L. A., & Sabil, S. (2019). The reliability and factor structure of the Westside Test Anxiety Scale among university students. *Online Journal of Health and Allied Sciences, 18*(3), 8.

Von der Embse, N., Jester, D., Roy, D., & Post, J. (2018). Test anxiety effects, predictors, and correlation: A 30-year meta-analytic review. *Journal of Affective Disorders, 227*, 483–493. https://doi.org/10.1016/j.jad.2017.11.048

Yuksel, M. Y., Sevim, E., & Çelimli, Ç. (2018). Examination of the relationship between test anxiety and selective attention among adolescents. *Elementary Education Online, 17*(2), 864–873. https://doi.org/10.17051/ilkonline.2018.419331