Psychometric Validation of Teacher Educators' Self-Assessment Test Scale for Quality and Equity in Spiritual Values Assessment

***Shetu Ilu Chetle, Adibah Abdul Latif**

Faculty of Educational Sciences and Technology, Universiti Teknologi Malaysia, Johor Bahru

DOI: https://doi.org/10.51584/IJRIAS.2025.10030062

Received: 09 March 2025; Accepted: 17 March 2025; Published: 19 April 2025

ABSTRACT

Psychometric validation of instruments measuring spiritual values in education is critical for ensuring fairness and quality in assessments within the Nigerian educational system. Spiritual values are integral to teacher educators, shaping their influence on students' moral and ethical development. This study aimed to validate the Teacher Educators Self-Assessment Test (TEDSAT) instrument using the Rasch Model for evaluating equity and quality in spiritual values assessment. The 25-item, four-point Likert-scale instrument underwent expert content validation, yielding strong content validity indices (I-CVI: 0.72–0.92). A pilot test with 37 teacher educators revealed moderate reliability (Cronbach's alpha: 0.70; item reliability: 0.77). Rasch analysis assessed item fit, unidimensionality, and difficulty levels. The findings showed that while the TEDSAT scale demonstrated overall validity and reliability, certain items required refinement for optimal performance. The study concludes that psychometric validation is essential for fair assessments of teacher educators' spiritual values, supporting quality educational practices in Nigeria.

Key Word: Spiritual values, Rasch Model Analysis, Equity and Quality Assessment

INTRODUCTION

The concept of spiritual values in education has gained recognition as a critical factor in holistic human development and ethical behavior. Spirituality, as a multidimensional construct, encompasses the search for meaning, transcendence, and higher purpose (Machul et al., 2022). It manifests in values such as compassion, care, integrity, respect, and altruism, which are essential for meaningful relationships and societal growth (Nazam, 2022; Nazam & Husain, 2020a). These values enable individuals to lead purpose-driven lives and positively impact their workplaces, communities, and broader society (McGhee & Grant, 2008). Within education, spirituality shapes teachers' moral and ethical practices, influencing student outcomes and fostering inclusive learning environments (Pandya, 2017; Husamah et al., 2022).

In Nigeria, the National Policy on Education (FRN, 2013) underscores the importance of moral and spiritual principles in quality education at all levels. The National Teacher Education Policy (FRN, 2014) advocates for training programs assessing educators' knowledge, skills, values, and attitudes. However, research indicates inconsistent implementation of spiritual values assessment (Oyedeji, 2015; FRN, 2022). Internationally, studies show spirituality enhances teacher well-being, job satisfaction, and ethical decision-

Comment [VN1]: The variable at the level of teacher educator being assessed, is missed out. Similarly in the abstract, the level is not missed . This makes the scope quite open/broad.

The values are five and the topic cannot contain them. Spiritual values is compound word or concept.

Comment [VN2]: Edit making, directly impacting students' academic and personal development (Lee, 2020; Mukherjee & Ghosh, 2022; Pong, 2022). Teachers who incorporate spiritual values create nurturing environments, fostering empathy, trust, and mutual respect (Nazam & Husain, 2020b; van Niekerk, 2014).

Despite global advancements, assessing spiritual values remains challenging due to their abstract and culturally sensitive nature (Giske et al., 2022; Robinson, 2017). This is particularly evident in Nigeria, where the lack of

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue III March 2025

validated instruments hinders efforts to ensure fairness and quality in educational evaluations (Akram et al., 2015; Rezai, 2022). Construct validation in educational assessments is crucial to ensure reliability and accuracy (Harmeni, 2022; Roy et al., 2023). Without proper validation, instruments risk producing unreliable and biased results, leading to misleading conclusions about teachers' spiritual capacities. Self-assessment tests, commonly used for measuring spiritual values, are prone to social desirability biases, further highlighting the need for rigorous psychometric evaluation (Hair et al., 2010; King et al., 2006).

The Rasch Model, a widely recognized psychometric framework, offers a robust approach to evaluating the reliability, validity, and fairness of assessment instruments by examining item fit, unidimensionality, and difficulty levels (Livingston, 2020; Ogunsakin & Shogbesan, 2018). In Nigeria, studies on assessing spiritual values among teacher educators remain limited, with most research focusing on student populations (Akram et al., 2015; Wang et al., 2020). This underscores the need for research addressing spiritual constructs relevant to teacher educators, including religious tolerance, trust, altruistic love, respect, and honesty (Nazam, 2022; Gui et al., 2020). Additionally, few studies apply advanced psychometric techniques like the Rasch Model to validate instruments measuring spiritual values (George, 2022; Vem et al., 2022).

To address these gaps, this study validates the Teacher Educators Self-Assessment Test (TEDSAT), a 25-item instrument measuring spiritual values using a four-point Likert scale. The Rasch Model evaluates its psychometric properties to ensure reliability and validity.

This research contributes to improving educational assessment practices in Nigeria, emphasizing spiritual values as critical components of equity assessment aligned with national policies and global educational development goals. Equity means that Assessment practices (Aps) should be adapted to test-takers' needs /characteristics (Rasooli et al., 2022; Rezai, 2022). According to Karlen et al., (2023); Monteiro et al., (2021), assessment practices are considered as fair if they do not unduly privilege a particular group of test-takers. Quality items are obtained only if educators first conduct an item analysis of the constructs before assessing learning outcomes (Azizah et al., 2022). Item analysis activities are conducted to review each question before use, enhance test item quality through revision, remove ineffective questions, and also find out the diagnostic information regarding the learner about the material that has been taught Rezigalla, 2022). Research shows that Rasch Model is can guide in this direction.

In the Rasch model, people are given the characteristics of the level of latent ability and the items are given the characteristics of the level of difficulty(Azizah et al., 2022). The probability of answering correctly is a function of the ratio between the level of ability and the difficulty of the item. An important feature of the one parameter (Rasch model) is that it does not contain discrimination and guess parameters (Rasch models for rating scale analysis, 2023; Karlin & Karlin, n.d.). In this model, it is assumed that item difficulty is the only item characteristic that affects test performance. In addition, in the Rasch model, the problem of selecting items to construct a test is an effort to plan a quality test according to the needs and objectives of the test (Krishnan & Idris, 2014; Saidi & Siew, 2019)**.** It is a method for determining the psychometric properties of a set of items or questions, and it allows decisions to be made about the composition and structure of a questionnaire; it can also be used to create interval measures for latent scales (Boone, 2016).

Table 1. Summary of item validity and reliability using Rasch Model

Criteria	**Statistical information**	**Results**
Item Validity	a. Item Polarity	PTMEA CORR > 0 (Bond & Fox 2007)
Item Fit	b. Item Fit	Total Mean Square infit and outfit of 0.6 – 1.4 (Bond & Fox 2007)
Item Misfit	c. Separation (SE)	All items show ≥ 2.0 (Linacre 2007)

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue III March 2025

Person Reliability	d. Person Reliability	Value > 0.8 (Bond & Fox 2007)
Item Reliability	e. Item Reliability	Value > 0.8 (Bond & Fox 2007)

Adopted from (Krishnan & Idris, 2014; Yasin et al., 2015).

This study used Principal Component Analysis (PCA) analysis in Rasch Model to check unidimensionality. The raw variance explained by the measures expected minimum percentage of 20%. The unexplained variance in the first contrast should have maximum 15% (Fisher, 2007 as cited in Rosli et al., 2020). The distribution of items and the ability of the subject to respond to items in the person-item map (Wright & Mok, 2004). The left side of the person-item map shows respondents' abilities, while the right side is the distribution of items (item difficulty). The easiest items are those in the lowest position, while the most difficult items are those at the top right (Azizah et al., 2022). The positive values in person-item map indicate values that are less familiar (difficult) to the respondents while the negative values are at the bottom side of the scale indicating the most familiar (easy) items to the respondents Wright, and (Mok, 2004; Boone, 2014). Thus, the need to determine the validity, reliability, item fit, unidimensionality and item difficulty level of TEDSAT spiritual value scale.

Research Objective: The objective of the study is to find out TEDSAT scale validity and reliability for optimal performance.

The research intents to answer the following research questions to enable researchers elicit information from respondents:

What is the reliability of TEDSAT scale?

What is the validity of TEDSAT scale for equity and quality assessment?

In terms of item fit statistics

unidimensionality, and difficulty levels

Comment [VN3]: The research objectives and questions are correctly aligned to the study focus however the disturbing questions is: which and whose standards are these activity items measure against? This questions also seeks to be addresse

LITERATURE REVIEW

The integration of spiritual values in education fosters holistic development and ethical behavior among teachers and students. Spiritual values involve the pursuit of meaning, self- discovery, and transcendence, reflected in principles such as respect, trust, integrity, and altruistic love (Nazam, 2022; Nazam & Husain, 2020a). These values create inclusive and compassionate educational environments, where teachers act as role models influencing students' moral and ethical growth (Pandya, 2017; Mukherjee & Ghosh, 2022).

Globally, spirituality in education is linked to teacher well-being, job satisfaction, and student outcomes, demonstrating its transformative impact (Husamah et al., 2022). Teachers who incorporate spiritual values foster empathy, compassion, and self-awareness, creating equitable learning environments (Nazam & Husain, 2020b). In Nigeria, the National Policy on Education (FRN, 2013) highlights the role of spiritual values in promoting quality education and interpersonal harmony. The National Teacher Education Policy (FRN, 2014) emphasizes assessing teachers' knowledge, skills, and values for holistic educator development. Despite these policies, research suggests inconsistent implementation of spiritual values assessment in teacher education (Oyedeji, 2015; FRN, 2022).

Assessing spiritual values poses challenges due to cultural variations (Giske et al., 2022). Self-assessment tests, widely used for evaluating spiritual values, help educators reflect on ethical practices (King et al., 2006; Axiak,

2023). However, these tools are prone to biases, including social desirability, affecting validity and reliability (Hair et al., 2010; Harmeni, at the broad study level focus on testing 'validity equity" of the testing model , the question is: against which/whose standards? Kindly check the table 1 before your comments/observation on page 3 for the correction. 2022). Psychometric validation is essential for ensuring that assessment tools accurately capture spiritual values (Rezai, 2022; Roy et al., 2023). Construct validation is particularly critical in spiritual values assessment, as inaccurate measurements can lead to misleading conclusions about educators' ethical capacities (Nazam & Husain, 2020b).

Studies highlight the importance of psychometric approaches, such as the Rasch Model, in validating self-assessment tools. The Rasch Model provides a robust framework for evaluating instrument reliability, validity, and fairness by analyzing item fit, unidimensionality, and difficulty levels (Livingston, 2020; Ogunsakin & Shogbesan, 2018). This approach ensures equity in assessment practices by accounting for respondent variations (Rezai, 2022).

Despite growing literature on spirituality in education, research on assessing spiritual values among Nigerian teacher educators is scarce. Most studies focus on student populations, neglecting teacher educators' roles in promoting spiritual values (Akram et al., 2015; Wang et al., 2020). Additionally, few studies use advanced psychometric techniques, such as the Rasch Model, to validate instruments measuring spiritual values (George, 2022; Vem et al., 2022). Cultural differences further complicate generalizing findings, emphasizing the need for localized research (FRN, 2022; Neal et al., 2022).

The Teacher Educators Self-Assessment Test (TEDSAT) was developed to address these gaps by providing a standardized tool for measuring spiritual values among Nigerian teacher educators. The instrument focuses on five key constructs—religious tolerance, trust, altruistic love, respect, and honesty essential for fostering ethical practices and positive relationships in educational settings (Nazam, 2022; Gui et al., 2020). Previous studies confirm these constructs' importance in creating supportive learning environments (Furqon Hidayatullah & Nurkamto, 2020; Pong, 2022).

The validation of TEDSAT using the Rasch Model enhances the accuracy and reliability of spiritual values assessment in Nigeria. Aligning with principles of fairness and equity, this study contributes to improving educational assessment practices and fostering ethical development among teacher educators.

METHODOLOGY INSTRUMENT DEVELOPMENT

The instrument used in this study, the Teacher Educators' Self-Assessment Test (TEDSAT), was self-developed based on a comprehensive review of the literature on spiritual value constructs. The TEDSAT instrument was designed to measure five key constructs: religious tolerance, trust, altruistic love, respect, and honesty. Each construct was represented by five items, resulting in a total of 25 items formatted as a four-point Likert scale (1 = Strongly Disagree to 4 = Strongly Agree). The primary purpose of the TEDSAT questionnaire was to enable teacher educators to self-assess their spiritual values as they relate to their interactions with student teachers, management staff, colleagues, and the broader college community. The focus was on promoting a safe and inclusive educational environment.

Content Validation

The TEDSAT instrument underwent rigorous content validation to ensure its relevance and appropriateness. A panel of six experts with experience in psychometrics and teacher education reviewed the 25 items for content accuracy, clarity, and alignment with the intended constructs. Minor revisions, such as rephrasing of items, were made based on the experts' feedback. The raters' evaluations were then analyzed to calculate the content validity indices for individual items (I-CVI) and the overall scale (S-CVI). Two approaches were used to determine the S-CVI: universal agreement (S-CVI/UA) and average approach (S- CVI/Ave). The results demonstrated good content validity, with indices ranging from 0.72 to 0.92, which are considered acceptable for instrument validation.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue III March 2025

Pilot Study

The TEDSAT instrument was subjected to a feasibility study through a pilot test following content validation. The sample for the pilot study comprised 37 teacher educators from Shehu Shagari College of Education, Sokoto, Nigeria. Participants were selected to ensure diversity in gender, age, and ethnic backgrounds, representing the heterogeneous makeup of teacher educators in Nigeria. The sample size of 37 was deemed adequate for a pilot study based on recommendations from Johanson and Brooks (2010), who suggested a minimum of 30 participants for scale development. Although Viechtbauer et al. (2015) recommend a larger sample size of 59 for clinical pilot studies to achieve a high confidence level, the current study focused on educational assessment, where smaller samples are often sufficient for preliminary evaluations.

Rationale for Pilot Study

The pilot study aimed to evaluate the feasibility and reliability of the TEDSAT instrument before its application in a larger sample. By using a diverse group of teacher educators, the pilot study provided an opportunity to identify any issues with item clarity, response consistency, or cultural sensitivity. The pilot results ensured the instrument was suitable for assessing spiritual values in teacher educators and aligned with the broader objectives of equity and fairness in educational assessments.

RESULTS AND DISCUSSION RELIABILITY

The reliability of the Teacher Educators' Self-Assessment Test (TEDSAT) instrument was evaluated using person reliability, item reliability, and separation indices. The results, as presented in Table 1, indicate that the overall reliability of the instrument is moderate. The Cronbach's alpha value of 0.70 meets the acceptable threshold for internal consistency (Hayati Ishak et al., 2018). The person reliability ranged from 0.62 to 0.77, reflecting moderate consistency in responses among teacher educators. Meanwhile, the item reliability was 0.77, suggesting that the quality of the test items was adequate for measuring spiritual values. However, the person and item separation indices were lower than recommended standards.

Table 2: TEDSAT spiritual value instrument Reliability

Domain	Total Item	Cronbach Alpha (α)	Person Reliability and Separation	Item Reliability and separation
Spiritual values	25	0.70	0.62: 1.28	0.77: 1.84

According to Linacre (2005), a separation index greater than 2.0 indicates a robust ability to distinguish between respondents with varying abilities, while Fisher (2005) suggests that item reliability values greater than 0.80 are desirable. In this study, the person separation index was 1.28, and the item separation index was 1.84, both falling short of these benchmarks. This suggests that some items may need modification or removal to enhance the instrument's discriminatory power and overall reliability. Nonetheless, the observed item reliability values ranging from 0.63 to 0.77 may still be considered reasonable for pilot studies (Fisher, 2007; Linacre, 2014).

The item reliability coefficient for TEDSAT's measurement reliability was further analyzed at the construct level, as shown in Table 2. The results reveal substantial variability in reliability and separation indices across the five spiritual value constructs. Religious tolerance and altruistic love were the most reliable constructs, with item reliability values of 0.78 and 0.84, respectively, and item separation indices of 1.89 and 2.29. In contrast, the constructs of trust, respect, respectively), with separation indices below the acceptable threshold of 2.0. The person reliability for individual constructs was also low, ranging from 0.15 to 0.56, which indicates a high level of inconsistency in respondents' answers (Linacre, 2009). This is contrary to the person reliability value of > 0.8 and item reliability value of > 0.8 (Bond and Fox 2007). Hence, some items were refined while some were eliminated especially those that misfit the Rasch model (Krishnan & Idris, 2014; Saidi & Siew, 2019). The result is low compare to the standard specified in table 1 above by (Krishnan & Idris, 2014; Yasin et al., 2015).

Comment [VN4]: How was this challenge addressed in the tool before use? This again raise the early uestions

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue III March 2025

on whose standards were

Table 3: Spiritual values construct by construct Reliability religious values being measured against? This

Construct	Total item	Reliability		Reliability	
		Person	separation	Item	Separation
Religious tolerance	5	.40	.82	.78	1.89
Trust	5	.15	.42	.45	.90
Altruistic love	5	.46	.93	.84	2.29
Respect	5	.40	.82	.63	1.32
Honest	5	.56	.42	.49	.97

The low person reliability values (0.15–0.56) and separation indices at the construct level may be attributed to several factors. These include respondent fatigue during the survey, negligence in completing items, missing data, and potential biases in responses (Dolnicar et al., 2016; Adhikari, 2021; Kaur, 2021). Additionally, the use of a relatively small sample size (n = 37) may have contributed to the variability in reliability estimates. Increasing the sample size by 10% in future studies, as suggested by Tseng and Sim (2021), could address this issue and improve the reliability of the instrument.

These results also highlight the need for item revision and potential elimination of poorly performing items. Constructs such as trust, respect, and honesty showed item reliability values below 0.50, suggesting that the quality of test items within these domains is suboptimal. Similar findings have been noted in previous research, where ambiguous or poorly worded items compromised the reliability of self-assessment instruments (Azizah et al., 2022; Hayati Ishak et al., 2018). Revising items to improve clarity and alignment with the intended constructs, particularly for low-performing domains, is essential to enhance the instrument's psychometric properties.

Fit Analysis

The item fit analysis indicated that most items aligned well with the Rasch Model's expectations. Items with acceptable infit and outfit mean square (MNSQ) values ranged between 0.6 and 1.4, as recommended by Boone et al. (2014). However, as detailed in Table variability with low indices below the threshold o these constructs: honesty, trust and respect rev the dilemma in standards being gauged against. T corroboration with longitudinal observation may address this dilemma, triangulation of instrumen2, two items—RT1 and RS5—were flagged for misfit. RT1 had a negative Point Measure Correlation (PTMEA) value of -0.04, indicating a lack of alignment between item responses and the overall construct. Similarly, RS5 showed a low PTMEA value of 0.25, suggesting that the item did not adequately distinguish between respondents' levels of spiritual values. These findings align with prior studies that emphasize the importance of reviewing and refining misfitting items to improve instrument validity (Kowiyah et al., 2020; Hayati Ishak et al., 2018).

Table 4: Infit and misfit

ITEM	Infit		Outfit		PT MEA CORR	CATEGORY
	MNSQ	ZSTD	MNSQ	ZSTD		
RT1					-.0.4	Fail
RS5	1.87	3.1	1.94	3.1	.25	Fair

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue III March 2025

Unidimensionality

The unidimensionality of the TEDSAT instrument was evaluated using Principal Component Analysis (PCA) of residuals, a common approach for assessing whether an instrument measures a single latent construct (Bond & Fox, 2007). This analysis ensures that the TEDSAT questionnaire captures the underlying construct of spiritual values without significant interference from other dimensions. Table 4 summarizes the results of the PCA.

Table 5: Principal Component Analysis (PCA)

Measures	Value
Raw Variance Explained by Measures	20.1%
Unexplained Variance in First Contrast	13.3%

The PCA results showed that the Raw Variance Explained by Measures was 20.1%, meeting the minimum threshold of 20% required to confirm unidimensionality (Bond & Fox, 2007). This indicates that the majority of the variance in the data can be attributed to the spiritual values construct. Furthermore, the Unexplained Variance in the First Contrast was 13.3%, which is below the recommended maximum of 15%. These results provide strong evidence that the TEDSAT instrument measures a single latent construct.

Unidimensionality is a critical psychometric property for ensuring the validity of an assessment tool. A unidimensional scale enables researchers to confidently interpret the test scores as a reflection of the intended construct—in this case, spiritual values. The results of the PCA align with previous studies that emphasize the importance of unidimensionality in constructing reliable and valid measurement instruments (Meijer & Tendeiro, 2017; Rezai, 2022).

These findings are significant in confirming the construct validity of the TEDSAT instrument. By demonstrating unidimensionality, the study ensures that the instrument is focused and does not confound the measurement of spiritual values with unrelated dimensions. This property is particularly important for ensuring the fairness and accuracy of assessments in educational settings, where clear and focused measurement is essential (Rasooli et al., 2023).

Difficulty Estimation

The item difficulty levels of the TEDSAT instrument were analyzed using the person-item map generated through the Rasch Model. The person-item map provides a visual representation of the distribution of respondent abilities (person logits) and item difficulties (item logits) on the same scale, allowing for an assessment of how well the items align with the spiritual value levels of the respondents. The map is shown in Figure 1.

PERSON - MAP – ITEM <more>|<rare>

2	+		
	|		
	|		
	|		
	|		
	|		
	|		

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue III March 2025

	|			
	X T|			
	|			
	X |			
	X |			
1	XX +			
	X S|			
	XX |T			
	XXX |			
	XXX | TR1	TR4		
	XXX | TR5			
	XXXXX M| RT4	RT5		
	XXXX |S			
	XXX |			
	X | RS1	RS5	TR2	
	| AL2	AL3	HO3	TR3
	X S| AL4	HO4		
0	XX +M			
	| RS3			
	| RT3			
	X | HO1			
	XX T| HO2	RS2	RT1	
	X |S HO5	RS4	RT2	
	|			
	|			
	| AL1			
	|			
	|T			
	|			
-1	+			
	| AL5			

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue III March 2025

	|			

-2 +

<less>|<frequ>

Figure 1: Spiritual Values of Teacher Educator's Person-Item Map

The person-item map reveals that the item difficulty levels ranged from -2.00 logits (simplest item) to +2.20 logits (most difficult item). The simplest item on the TEDSAT scale was identified as AL5, representing the altruistic love construct. Conversely, the most difficult items, TR1 and TR4, were associated with the trust construct. These results demonstrate a logical spread of item difficulty, with the majority of items falling within the moderate range of 0.00 to +1.00 logits, aligning well with the ability levels of most respondents. This aligns with Equity meaning that Assessment practices (Aps) should be adapted to test-takers' needs/characteristics (Rasooli et al., 2022; Rezai, 2022). According to Karlen et al., (2023); Monteiro et al., (2021), assessment practices are considered as fair if they do not unduly privilege a particular group of test-takers

However, an imbalance was observed in the coverage of higher difficulty levels. The map indicates that there were relatively few items targeted at respondents with higher abilities (+2 logits or more). This gap suggests that the TEDSAT instrument may not adequately measure the spiritual values of individuals with exceptionally high levels of spiritual awareness. This finding aligns with Wright & Stone, (1999); Bond and Fox (2007), who emphasize the importance of item coverage across the entire range of respondent abilities in educational assessments.

The results also highlight specific constructs where item difficulty varied significantly. The trust construct, despite containing the most challenging items (TR1 and TR4), demonstrated a relatively low overall alignment with respondent abilities. This could be due to poorly worded or ambiguous items, which might fail to differentiate effectively between respondents with varying levels of trust (Hayati Ishak et al., 2018; Darmana et al., 2021).

The alignment of item difficulties with respondent abilities is a critical aspect of instrument validity. A well-aligned scale ensures that respondents are neither over-challenged by excessively difficult items nor under-assessed by items that are too easy. Previous studies have highlighted the importance of ensuring a balance in item difficulty to accurately capture the full range of latent traits (Wright & Stone, 1999; Futri et al., 2022).

Summary

This study aimed to validate the psychometric properties of the Teacher Educators' Self- Assessment Test (TEDSAT) instrument, focusing on its reliability, validity, and unidimensionality. The major findings are summarized as follows:

Reliability Analysis

The overall reliability of the TEDSAT instrument was moderate, with a Cronbach's alpha of 0.70, item reliability of 0.77, and person reliability ranging from 0.62 to 0.77. While these values meet acceptable thresholds for internal consistency, the separation indices for person (1.28) and item (1.84) fell below the recommended benchmark of >2.0 (Linacre, 2005; Fisher, 2007). At the construct level, reliability varied, with Religious Tolerance (0.78) and Altruistic Love (0.84) demonstrating acceptable item reliability, while Trust (0.45), Respect (0.63), and Honesty (0.49) exhibited lower reliability. These results suggest that some items within the latter constructs may require refinement or removal to enhance their quality (Azizah et al., 2022; Darmana et al., 2021).

Unidimensionality

The Principal Component Analysis (PCA) of residuals confirmed the unidimensionality of the TEDSAT instrument. The Raw Variance Explained by Measures was 20.1%, exceeding the minimum threshold of 20% (Bond & Fox, 2007). Furthermore, the Unexplained Variance in the First Contrast was 13.3%, below the

recommended maximum of 15%, indicating that the instrument measures a single latent construct spiritual values. This finding supports the validity of the TEDSAT scale and its alignment with the theoretical framework underpinning spiritual values assessment (Meijer & Tendeiro, 2017).

Item Difficulty

The person-item map analysis revealed that the TEDSAT items were generally well- aligned with respondent ability levels, with most items falling within the moderate difficulty range of 0 to +1 logits. However, gaps were identified in the higher ability range, where there were insufficient items to assess respondents with high levels of spiritual values. This limitation was particularly evident in the Trust construct, which contained the most difficult items but showed poor alignment with respondent abilities. These findings highlight the need to develop additional challenging items to ensure comprehensive coverage of the latent trait (Wright & Stone, 1999; Bond & Fox, 2007).

Misfitting Items

Two items, RT1 (Religious Tolerance) and RS5 (Respect), were identified as misfitting based on their Point Measure Correlation (PTMEA) values of -0.04 and 0.25, respectively. These results suggest that these items failed to align with the overall construct of spiritual values. Misfitting items can negatively impact the reliability and validity of an instrument and should be revised or removed in future iterations to improve overall measurement quality (Kowiyah et al., 2020; Hayati Ishak et al., 2018).

Construct Performance

Among the five spiritual value constructs, Altruistic Love and Religious Tolerance exhibited strong psychometric properties, with high item reliability and separation indices. Conversely, Trust, Respect, and Honesty showed lower reliability and separation values, indicating the need for item-level refinements. This variability across constructs underscores the importance of conducting construct-specific analyses to identify areas for improvement (Futri et al., 2022; Darmana et al., 2021).

CONCLUSION AND RECOMMENDATIONS

This study validated the Teacher Educators' Self-Assessment Test (TEDSAT) as a psychometric tool designed to measure spiritual values among teacher educators, focusing on constructs such as Religious Tolerance, Trust, Altruistic Love, Respect, and Honesty. The findings demonstrated that the TEDSAT instrument has moderate reliability, unidimensionality, and alignment with the Rasch Model, making it a valuable resource for assessing spiritual values in educational settings. However, specific areas, such as low item reliability for certain constructs, misfitting items, and insufficient representation of higher ability levels, require refinement to enhance the instrument's validity and overall utility for usage in the actual research for spiritual value assessment in Nigerian colleges of Education and in assessment of value education in the country's educational system and beyond. Further researchers are spur to adopt longitudinal observational tools in determining the spiritual values in their assessments.

The study confirmed that the TEDSAT scale measures a single latent construct, as evidenced by the Principal Component Analysis of residuals, which demonstrated acceptable levels of explained variance and minimal unexplained variance. This unidimensionality supports the instrument's theoretical framework and provides a solid foundation for its application in research and practice. Despite these strengths, the study identified limitations, including low reliability and separation indices for certain constructs and a lack of challenging items for respondents with higher levels of spiritual values. Addressing these issues is critical to improving the instrument's psychometric properties and ensuring its effectiveness in diverse educational contexts.

Based on the findings, the following recommendations are proposed for future development and application of the TEDSAT instrument:

Refinement of Misfitting Items:

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue III March 2025

Items such as RT1 (Religious Tolerance) and RS5 (Respect), which demonstrated poor fit and low point-measure correlations, should be revised or eliminated. Item revision should focus on ensuring clarity, cultural relevance, and alignment with the intended constructs to enhance their discriminatory power.

Enhancing Construct Reliability:

Constructs with low item reliability, such as Trust (0.45) and Honesty (0.49), require additional refinement. Developing new items that align with these constructs' theoretical underpinnings will improve their reliability and contribute to a more balanced scale.

Addressing Gaps in Item Difficulty:

The lack of items targeting higher ability levels (+2 logits or above) necessitates the development of additional challenging items. Expanding the difficulty range will ensure that the TEDSAT instrument can comprehensively assess spiritual values across a wider spectrum of respondent abilities.

Increasing Sample Size:

The sample size of 37 participants, though sufficient for a pilot study, may have contributed to variability in reliability and consistency. Increasing the sample size by at least 10%, as recommended by Tseng and Sim (2021), would provide more robust data and reduce the impact of response inconsistencies.

Continuous Validation:

To strengthen the validity of the TEDSAT instrument, ongoing validation studies should be conducted using larger and more diverse samples. These studies should explore the instrument's performance across different cultural, institutional, and demographic contexts to ensure its generalizability.

Implications

The validation of the TEDSAT instrument has significant implications for research, practice, and policy in education.

Theoretical Implications:

This study contributes to the growing body of literature on spiritual values in education by providing a validated framework for assessing constructs such as Religious Tolerance, Trust, and Altruistic Love. By confirming the unidimensionality of the TEDSAT scale, the study establishes its theoretical grounding and highlights the importance of spiritual values as a distinct construct in teacher education research.

Practical Implications:

The TEDSAT instrument offers a valuable tool for teacher education programs seeking to promote ethical and spiritual values among educators. Institutions can use the instrument for self-assessment, professional development, and curriculum design, aligning with broader goals of fostering moral and ethical behavior in schools and colleges. Additionally, the identification of areas for improvement provides actionable insights for enhancing the instrument’s reliability and applicability in diverse educational settings but first in the main research (thesis) to determine its usability.

Policy Implications:

The findings align with Nigeria's National Policy on Education (FRN, 2013), which emphasizes the integration of spiritual and moral values in teacher education. By providing a validated instrument for measuring spiritual values, this study supports the development of evidence-based policies that prioritize equity, fairness, and ethical leadership in education. Policymakers can leverage the TEDSAT instrument to monitor and evaluate the implementation of spiritual and moral education programs, ensuring alignment with national goals and standards.

REFERENCES

1. Abdaziz, A., Jusoh, M., Amlus, H., & SALWANI Salleh, T. (2014). Construct Validity: A Rasch Measurement Model Approaches The Influences of Attitude Towards Behaviour, Subjective Norms and Self-Efficacy on Entreprenuerial Intention among the Technical Secondary Students in Malaysia View project. https://www.researchgate.net/publication/266676182
2. Akram, M., Zepeda, S., & Zepeda, S. J. (2015). Development and Validation of a Teacher Self-assessment Instrument. In Journal of Research and Reflections (Vol. 9, Issue 2). http://www.ue.edu.pk/jrre
3. Ali Rezigalla, A. (2022). Item Analysis: Concept and Application. In Medical Education for the 21st Century. IntechOpen. https://doi.org/10.5772/intechopen.100138
4. A Tool for Performing Item Analysis To Enhance Teaching And Learning Experiences. (2017). Issues In Information Systems. https://doi.org/10.48009/1_iis_2017_128-136
5. Axiak, G. (2023). Self-Assessment-Its Good and Bad Points. https://www.researchgate.net/publication/371082082
6. Azizah, N., Suseno, M., & Hayat, B. (2022). Item analysis of the rasch model items in the final semester exam indonesian language lesson. World Journal of English Language, 12(1), 15–26. https://doi.org/10.5430/wjel.v12n1p15
7. Bond, T. G. (2010). Rasch Model. In The Corsini Encyclopedia of Psychology (pp. 1–1). Wiley. https://doi.org/10.1002/9780470479216.corpsy0773
8. Bond, T. G., & Fox, C. M. (2007a). Applying the Rasch Model: Fundamental Measurement in the Human Sciences Second Edition. www.erlbaum.coOl
9. Bond, T. G., & Fox, C. M. (2007b). Applying the Rasch Model: Fundamental Measurement in the Human Sciences Second Edition. www.erlbaum.coOl
10. Boone, W. J. (2016). Rasch analysis for instrument development: Why,when,and how? CBE Life Sciences Education, 15(4). https://doi.org/10.1187/cbe.16-04-0148
11. Boone, W. J., Staver, J. R., & Yale, M. S. (2014a). Person Reliability, Item Reliability, and More. In Rasch Analysis in the Human Sciences (pp. 217–234). Springer Netherlands. https://doi.org/10.1007/978-94-007-6857-4_10
12. Boone, W. J., Staver, J. R., & Yale, M. S. (2014b). Person Reliability, Item Reliability, and More. In Rasch Analysis in the Human Sciences (pp. 217–234). Springer Netherlands. https://doi.org/10.1007/978-94-007-6857-4_10
13. Braam Rust, A. A., & Gabriels, C. E. C. (2018). Spirituality in the workplace: Awareness of the human resources function. In International Journal of Accounting, Auditing and Taxation (Vol. 5, Issue 4). www.internationalscholarsjournals.org
14. Broadbent, J., Sharman, S., Panadero, E., & Fuller-Tyszkiewicz, M. (2021). How does self- regulated learning influence formative assessment and summative grade? Comparing online and blended learners. Internet and Higher Education, 50. https://doi.org/10.1016/j.iheduc.2021.100805
15. Dai, S., Vo, T. T., Kehinde, O. J., He, H., Xue, Y., Demir, C., & Wang, X. (2021). Performance of Polytomous IRT Models With Rating Scale Data: An Investigation Over Sample Size, Instrument Length, and Missing Data. Frontiers in Education, 6. https://doi.org/10.3389/feduc.2021.721963
16. Darmana, A., Sutiani, A., Nasution, H. A., Ismanisa*, I., & Nurhaswinda, N. (2021). Analysis of Rasch Model for the Validation of Chemistry National Exam Instruments. Jurnal Pendidikan Sains Indonesia, 9(3), 329–345. https://doi.org/10.24815/jpsi.v9i3.19618
17. DeSantis, A. J. (2020). Kohlberg's Moral Development Theory and Evolutionary Morality Model on Views of Homosexuality. Human Ethology, 35(1), 106–121. https://doi.org/10.22330/he/35/106-121
18. Dolnicar, S., Grün, B., & Leisch, F. (2016). Increasing sample size compensates for data problems in segmentation studies. Journal of Business Research, 69(2), 992–999. https://doi.org/10.1016/j.jbusres.2015.09.004
19. FRN. (2013). National Education Policy-2013_013213.
20. FRN. (2022). Manual of Critical Assessment Competences for Tertiary Education Lecturers in Nigeria.
21. FRN National Teacher Education Policy. (2014). Federal Republic of Nigeria.
22. Furqon Hidayatullah, M., & Nurkamto, J. (2020). Spirituality in Education: Perception of Teachers of

Middle School Teachers in Sukoharjo District.
23. Futri, V. I., Rosnawati, R., Rahim, A., & Marlina, M. (2022). Rasch Model Study on Mathematics Examination Test Using Item Response Theory Approach. International Journal on Emerging Mathematics Education, 6(1), 29. https://doi.org/10.12928/ijeme.v6i1.21761
24. George, J. A. (2022). Assessing Affective Domain in Higher Education: Educator's Perception. https://doi.org/10.17605/OSF.IO/ASD49
25. Giske, T., Schep-Akkerman, A., Bø, B., Cone, P. H., Moene Kuven, B., Mcsherry, W., Owusu, B., Ueland, V., Lassche-Scheffer, J., van Leeuwen, R., & Ross, L. (2023). Developing and testing the EPICC Spiritual Care Competency Self-Assessment Tool for student nurses and midwives. Journal of Clinical Nursing, 32(7–8), 1148–1162. https://doi.org/10.1111/jocn.16261
26. Grigoropoulos, J. E. (2019). The Role of Ethics in 21st Century Organization. International Journal of Progressive Education, 15(2), 167–175. https://doi.org/10.29329/ijpe.2019.189.12
27. Grosemans, I., Smet, K., Houben, E., De Cuyper, N., & Kyndt, E. (2020). Development and validation of an instrument to measure work-related learning. Scandinavian Journal of Work and Organizational Psychology, 5(1). https://doi.org/10.16993/SJWOP.99
28. Gui, A. K. W., Yasin, M., Abdullah, N. S. M., & Saharuddin, N. (2020). Roles of teacher and challenges in developing students' morality. Universal Journal of Educational Research, 8(3 3C), 52–59. https://doi.org/10.13189/ujer.2020.081606
29. Harmeni, H. A. & T. R. (2022). Construct validity analysis: Assessment for learning primary mathematics questionnaire. Journal for Educators, Teachers and Trainers, 13(2). https://doi.org/10.47750/jett.2022.13.02.006
30. Hayati Ishak, A., Rahimi Osman, M., Hapiz Mahaiyadin, M., Amzari Tumiran, M., & Anas, N. (2018). Article ID: IJCIET_09_09_141-Available online at Cite this Article: Amal Hayati Ishak, Muhamad Rahimi Osman, Mohd Hapiz Mahaiyadin, Mohd Amzari Tumiran and Norazmi Anas Examining , Unidimensionality of Psychometric Properties Via Rasch Model. In International Journal of Civil Engineering and Technology (IJCIET (Vol. 9, Issue 9). http://iaeme.com/Home/journal/IJCIET1462editor@iaeme.comhttp://iaeme.comhttp://iaeme. com/Home/journal/IJCIET1463http://iaeme.com
31. Hemachandran, A., Patwary, D., Dovina, D., & Professor, A. (2018). Relationship between Spirituality and Morality among Young Adults. In International Journal of Creative Research Thoughts (Vol. 6, Issue 1). www.ijcrt.orgwww.ijcrt.org
32. Huber, S. G., & Skedsmo, G. (2017). Standardization and assessment practices. Educational Assessment, Evaluation and Accountability, 29(1), 1–4. https://doi.org/10.1007/s11092-017- 9257-1
33. Husamah, H., Suwono, H., Nur, H., & Dharmawan, A. (2022). The development and validation of environmental literacy instrument based on spirituality for prospective science teachers. Eurasia Journal of Mathematics, Science and Technology Education, 18(12). https://doi.org/10.29333/EJMSTE/12732
34. Johanson, G. A., & Brooks, G. P. (2010). Initial scale development: Sample size for pilot studies. Educational and Psychological Measurement, 70(3), 394–400. https://doi.org/10.1177/0013164409355692
35. Karlen, Y., Bäuerlein, K., & Brunner, S. (2023). Teachers' assessment of self-regulated learning: Linking professional competences, assessment practices, and judgment accuracy. Social Psychology of Education. https://doi.org/10.1007/s11218-023-09845-4
36. Karlin, O., & Karlin, S. (n.d.). Making Better Tests with the Rasch Measurement Model (Vol. 13).
37. Kılıçoğlu, G. (2018). Study on the relationship between social studies course self-efficacy and motivation levels of secondary school students. Universal Journal of Educational Research, 6(8), 1743–1748. https://doi.org/10.13189/ujer.2018.060816
38. King, M. B., Jones, L., Low, J., & Walker, C. (2006). Measuring spiritual belief : development and standardization of a Beliefs and Values Scale. April. https://doi.org/10.1017/S003329170500629XKohlberg's Theory of Moral Development. (n.d.).
39. Kowiyah, K., Marini, A., & Wihardjo, S. (2020, January 17). Rasch Model Analysis of Critical Thinking Instruments for Elementary School. https://doi.org/10.4108/eai.30-9-2019.2291176
40. Krishnan, S., & Idris, N. (2014). Investigating Reliability and Validity for the Construct of Inferential Statistics. In International Journal of Learning, Teaching and Educational Research (Vol. 4, Issue 1).
41. Lee, J. C. (2020). Children ' s spirituality , life and values education : cultural , spiritual and educational

perspectives. https://doi.org/10.1080/1364436X.2020.1790774
42. Linacre, J. M. (2009). Requests for reprints should be sent to Local Independence and Residual Covariance: A Study of Olympic Figure Skating Ratings.
43. Livingston, S. A. (2020a). Basic Concepts of Item Response Theory: A Nonmathematical Introduction.
44. Livingston, S. A. (2020b). Basic Concepts of Item Response Theory: A Nonmathematical Introduction.
45. Machul, M., van Leeuwen, R., Ozga, D., Jurek, K., Boczkowska, S., & Dobrowolska, B. (2022). The level of spiritual care competence of Polish nurses and the psychometric properties of the spiritual care competence scale (SCCS). BMC Nursing, 21(1). https://doi.org/10.1186/s12912-022-00889-z
46. Makkar, S., & Singh, A. K. (2018). Development of a spirituality measurement scale.
47. Matazu S.S Julius. (2021). Item Analysis: A Veritable Tool for Effective Assessment in Teaching and Learning. Journal of Education and Practice. https://doi.org/10.7176/jep/12- 21-04
48. Mcghee, P., & Grant, P. (2008). Spirituality and Ethical Behaviour in the Workplace: Wishful Thinking or Authentic Reality. In Electronic Journal of Business Ethics and Organization Studies (Vol. 13, Issue 2). http://ejbo.jyu.fi/
49. McGhee, P., & Grant, P. (2017). The transcendent influence of spirituality on ethical action in organizations. Journal of Management, Spirituality and Religion, 14(2), 160–178. https://doi.org/10.1080/14766086.2016.1268539
50. Meijer, R. R., & Tendeiro, J. N. (2017). Unidimensional item response theory. In The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale and Test Development (Vols. 1–2, pp. 413–443). Wiley Blackwell. https://doi.org/10.1002/9781118489772.ch15
51. Monteiro, V., Mata, L., & Santos, N. N. (2021). Assessment Conceptions and Practices: Perspectives of Primary School Teachers and Students. Frontiers in Education, 6(April), 1–15. https://doi.org/10.3389/feduc.2021.631185
52. Mukherjee, S., & Ghosh, S. (2022). Enhancing Effectiveness of Values and Ethics Education through Spirituality. Journal of Management, Spirituality and Religion, 19(3), 275–301. https://doi.org/10.51327/CHXM7253
53. Nazam, F. (2022). Exploring Spiritual Values among School Children Exploring Spiritual Values among School Children. January 2016. https://doi.org/10.4172/2469-9837.1000175
54. Nazam, F., & Husain, A. (2020a). Further validation of spiritual values scale. New Educational Review, 61, 199–210. https://doi.org/10.15804/tner.2020.61.3.16
55. Nazam, F., & Husain, A. (2020b). Further validation of spiritual values scale. New Educational Review, 61, 199–210. https://doi.org/10.15804/tner.2020.61.3.16
56. Neal, J., Altman, Y., & Mayrhofer, W. (2022). Workplace spirituality: Making a difference. In Workplace Spirituality: Making a Difference. De Gruyter. https://doi.org/10.1515/9783110711349
57. O'Grady, E., Hinchion, C., & McNamara, P. M. (2011). The importance of respect in teaching and learning: Perspectives of final year pre-service teachers in a regional university in Ireland. European Journal of Teacher Education, 34(4), 501–518. https://doi.org/10.1080/02619768.2011.592978
58. Ogunsakin, I. B., & Shogbesan, Y. O. (2018). Item Response Theory (IRT): A Modern Statistical Theory for Solving Measurement Problem in 21st Century. In International Journal of Scientific Research in Education (Vol. 11, Issue 3B). http://www.ijsre.com.
59. Panadero, E. (2023). Toward a paradigm shift in feedback research: Five further steps influenced by self-regulated learning theory. Educational Psychologist, 58(3), 193–204. https://doi.org/10.1080/00461520.2023.2223642
60. Pandya, S. P. (2017). Spirituality and values education in elementary school: Understanding views of teachers. Children and Schools, 39(1), 33–41. https://doi.org/10.1093/cs/cdw042
61. Pong, H. K. (2022). The Correlation between Spiritual Well-Being and Burnout of Teachers. Religions, 13(8). https://doi.org/10.3390/rel13080760
62. Rasooli, A., Rasegh, A., & Firoozi, T. (2022). Conceptions of Fairness in Classroom Assessment: An Empirical Study Article. Journal of Teacher Education. https://doi.org/10.1177/00224871221130
63. Rasooli, A., Rasegh, A., Zandi, H., & Firoozi, T. (2023). Teachers' Conceptions of Fairness in Classroom Assessment: An Empirical Study. Journal of Teacher Education, 74(3), 260–273. https://doi.org/10.1177/00224871221130742
64. Rezai, A. (2022). Fairness in classroom assessment: development and validation of a questionnaire.

Language Testing in Asia, 12(1). https://doi.org/10.1186/s40468-022-00162-9
65. Robinson, C. (2017). Educators' practices for promoting the spiritual development of children Educators' practices for promoting the spiritual development of children aged 3 to 4 years, in the context of Catholic childcare centres in Western aged 3 to 4 years, in the context of Catholic childcare centres in Western Australia Australia. https://researchonline.nd.edu.au/theses/159
66. Rodríguez-Guevara, C., Mera-Mamián, A. Y., Muñoz-Rodríguez, D. I., Montenegro-Martínez, G., Roa-Urrutia, P. E., & Giraldo-Gallo, E. A. (2024). Validity and reliability of instruments for global measurement of disability: Systematic literature review. Salud UIS, 56(1). https://doi.org/10.18273/saluduis.56.e:24048
67. Roy, R., Sukumar, G. M., Philip, M., & Gopalakrishna, G. (2023). Face, content, criterion and construct validity assessment of a newly developed tool to assess and classify work–related stress (TAWS– 16). PLoS ONE, 18(1 January). https://doi.org/10.1371/journal.pone.0280189
68. Saidi, S. S., & Siew, N. M. (2019). Reliability and Validity Analysis of Statistical Reasoning Test Survey Instrument using the Rasch Measurement Model. International Electronic Journal of Mathematics Education, 14(3). https://doi.org/10.29333/iejme/5755
69. Scholarworks, S., & Georgeson-Agege, S. B. (n.d.). Conceptualising, Implementing, and Teaching Ethical Leadership Conceptualising, Implementing, and Teaching Ethical Leadership for Nigerian Public Officials for Nigerian Public Officials. https://scholarworks.waldenu.edu/dissertations
70. Semilarski, H., Soobard, R., & Rannikmäe, M. (2019). Modeling Students' Perceived Self- efficacy and Importance toward Core Ideas and Work and Life Skills in Science Education. Science Education International, 30(4), 261–273. https://doi.org/10.33828/sei.v30.i4.3
71. Siew, N. M., Syafiq, M., & Rahman, A. (2019). Assessing the Validity and Reliability of the Future Thinking Test using Rasch Measurement Model. In INTERNATIONAL JOURNAL OF ENVIRONMENTAL & SCIENCE EDUCATION (Vol. 14, Issue 4). http://www.ijese.com
72. Sovey, S., Osman, K., & Matore, M. E. E. M. (2022). Rasch Analysis for Disposition Levels of Computational Thinking Instrument Among Secondary School Students. Eurasia Journal of Mathematics, Science and Technology Education, 18(3), 2–15. https://doi.org/10.29333/ejmste/11794
73. State, C. R. (2020). Social value orientation and self-efficacy among secondary school students in. January 2019.
74. Subheesh, N. P., & Sethy, S. S. (2020). Learning through assessment and feedback practices: A critical review of engineering education settings. Eurasia Journal of Mathematics, Science and Technology Education, 16(3), 1–18. https://doi.org/10.29333/ejmste/114157
75. Suriadi, S., & Mursidin, M. (2019). Internalization of Cultural Values in Schools Early Steps in Building Nation Characters. 20. https://doi.org/10.4108/eai.8-12-2018.2283966
76. Tesio, L., Caronni, A., Kumbhare, D., & Scarano, S. (2024). Interpreting results from Rasch analysis 1. The ―most likely‖ measures coming from the model. Disability and Rehabilitation, 46(3), 591–603. https://doi.org/10.1080/09638288.2023.2169771TRCN FRN. (2013). CODE OF CONDUCT.
77. van Niekerk, P. (2014). Teachers and moral development of learners: A case for a Kohlbergian and neuroscience approach. Mediterranean Journal of Social Sciences, 5(4), 340–347. https://doi.org/10.5901/mjss.2014.v5n4p340
78. Vem, L. J., Tuamyil, R. B., Ocholi, S. A., Thurasamy, R. A. L., & Vem, L. J. (2022). Turnover Intention in Nigerian Universities: Do Academics' Spirituality and the Spiritual Climate Turnover Intention in Nigerian Universities: Do Academics' Spirituality and the Spiritual Climate Matter? May. https://doi.org/10.1080/18146627.2022.2063750
79. Viechtbauer, W., Smits, L., Kotz, D., Budé, L., Spigt, M., Serroyen, J., & Crutzen, R. (2015). A simple formula for the calculation of sample size in pilot studies. Journal of Clinical Epidemiology, 68(11), 1375–1379. https://doi.org/10.1016/j.jclinepi.2015.04.014
80. Wang, C.-F., Huang, S. S., & Lin, Y.-C. (2020). Construction and Validation of a Teacher Self-evaluation Scale to Measure Teaching Performance in a Medical University. Article in Journal of Medical Education. https://doi.org/10.6145/jme.202012_24(4).0001
81. Westbrook, C. J., Davis, D. E., McElroy, S. E., Brubaker, K., Choe, E., Karaga, S., Dooley, M., O'Bryant, B. L., Van Tongeren, D. R., & Hook, J. (2018). Trait Sources of Spirituality Scale: Assessing Trait Spirituality More Inclusively. Measurement and Evaluation in Counseling and Development, 51(2), 125–138. https://doi.org/10.1080/07481756.2017.1358059

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume X Issue III March 2025

82. Widodo, S., & Suryosukmono, G. (2021). Spiritual leadership, workplace spirituality and their effects on meaningful work: Self-transcendence as mediator role. Management Science Letters, 2115–2126. https://doi.org/10.5267/j.msl.2021.2.016
83. Widodo, W., Nyoman, I., Degeng, S., Pali, M., & Hitipeuw, I. (2024). Development and Validation of the Student Well-Being Scale for Elementary School Using Rasch Model Analysis. Journal of Educational, Health and Community Psychology, 13(2).
84. Wright, B., & Stone, M. (1999). Measurement Essentials 2nd Ed.
85. Yasin, R. Mohd., Yunus, F. A. N., Rus, R. C., Ahmad, A., & Rahim, M. B. (2015). Validity and Reliability Learning Transfer Item Using Rasch Measurement Model. Procedia - Social and Behavioral Sciences, 204, 212–217. https://doi.org/10.1016/j.sbspro.2015.08.143
86. Žujović, A. M. (2018). Predictive Validity of Florida' s Postsecondary Education Readiness Test.April.