AJEMATES
African Journal of Educational Management, Teaching and Entrepreneurship Studies
VOL.18 (1), May-August, 2026; https://ajemates.org

DEVELOPMENT AND STANDARDIZATION OF TEACHERS' JOB SATISFACTION SCALE: A RASCH MODEL-BASED APPROACH

Ifeoma Clementina Metu
Department of Educational Foundations,
Nnamdi Azikiwe University, Awka, Nigeria.
ic.metu@unizik.edu.ng
https://orcid.org/0000-0003-2968-2853

Barnabas Chidi Madu
Department of Science Education,
University of Nigeria, Nsukka- Enugu State, Nigeria.
barnabas.madu@unn.edu.ng

Chukwuemeka Aloysius Metu
Department of Curriculum and Instruction,
The University of Alabama, Tuscaloosa, USA
cametu@crimson.ua.edu
https://orcid.org/0009-0000-5185-133X

Corresponding author: ic.metu@unizik.edu.ng
08029773661

ABSTRACT

This research sought to develop and standardize a scale for teachers' job satisfaction utilizing the Rasch (IRT) model. This was in contrast to other researchers who made use of Classical Test Theory. Instrumentation and survey research design were utilized in this study, and a 90-item instrument called the Teachers' Job Satisfaction Scale (TJSS) was developed and trial tested. 55 items remained after factorial validation. The sample comprised 972 instructors from 36 schools in Nigeria, and data analysis was conducted using the Conditional Maximum Likelihood Estimation approach of WINSTEPS 5.1.1 software. 52 of the 55 items conform to the Rasch model, with item threshold parameter estimations exhibiting standard errors of measurement (SEM) below 0.5, indicating exact measurement of parameters. Item and person separation indices (10.75, 2.39), as well as reliability indices of 0.85 and 0.89 respectively showed that items had good alignment on the hierarchy and individuals had been divided into higher and lower levels of job satisfaction - with no gaps on the person/item map - meaning that all levels of the trait were covered by items on TJSS. It is recommended that the TJSS has powerful psychometric properties for measuring teachers' job satisfaction.

Keywords: Teachers, Job satisfaction, Rasch model, Threshold parameter, Separation indices, Unidimensionality

AJEMATES
African Journal of Educational Management, Teaching and Entrepreneurship Studies
VOL.18 (1), May-August, 2026; https://ajemates.org

Introduction

People work for a variety of reasons, including the desire for funds, activity, a sense of worth, respect from others, and social interaction. Personal demands, including robust self-esteem, the cultivation of amicable friendships, the enhancement of personal and geographical autonomy, and establishing a pleasant identity, constitute essential motivators for individuals' efforts. As a result, labour delivers physiological, social, and economic outputs that are essential for societal involvement. Work, for example, determines an individual's standing within a community, the individual's lifestyle, the family life and standards. When employees' social, physiological, and economic requirements are met effectively, they experience job satisfaction.

Job satisfaction was explained by Chaudry et al., (2024) as a pleasurable emotional reaction to job contents, however Cheng et al., (2022) claimed that job satisfaction is merely employees' psychological inclination towards their work. Locke (1976) sees job satisfaction as a favorable mental state derived from assessing one's occupation as fulfilling or contributing to one's job success. Locke posits that job dissatisfaction is an adverse emotional condition that results from perceiving an individual's employment as bothersome or obstructive to the realization of one's job values. Job satisfaction is defined as the degree to which a worker's fundamental physiological and psychological needs are fulfilled in their role, or the contentment derived from recognition and appreciation of one's contributions by others (Metu, 2020). It has frequently been asserted (Pepe, Addimando, &Veronesee, 2017; Zhou et al., 2024) that teachers' satisfaction with their jobs significantly impacts students' learning.

The role of teachers in the system of education is to assist students in developing knowledge, competencies, skills, and ethical character. Consequently, the contributions of educators are profoundly significant to others, both presently and in the future. A teacher is one who has undergone the required training to acquire the skills and knowledge, and accredited by an appropriate body (Federal Republic of Nigeria, 2013). Teachers are the bedrock of any nation, cultivating individuals from various disciplines. A cadre of competent and satisfied educators is essential for the educational system to achieve its intended objectives. Given that work satisfaction is positively connected with job performance and productivity, it is essential to develop an objective scale to assess teacher job satisfaction.

The development of instruments is a critical stage in the assessment and evaluation of education. Gall, Borg, and Gall (2015) established a set of criteria to assess the adequacy of a test or scale for use in educational research during instrument development. They consist of Objectivity: This indicates whether test findings are devoid of biases from the individuals who give and evaluate the test. Rasch (1980) posited that the metric employed to characterize an attribute must be independent of the observer to provide objectivity, a fundamental requirement for scientific measurement. Validity and Reliability: An effective measure

accurately assesses its intended construct, while a reliable measure consistently evaluates that construct across time, individuals, and contexts.

Measurement Theories

The two main measurement theories used by researchers to develop measures of latent characteristics are Classical Test Theory (CTT) and Item Response Theory (IRT). Latent characteristics, which are unobservable by definition, must be quantified indirectly using a test, task, or survey. Unobservable characteristics can be analyzed in this manner since it is assumed that they influence how respondents respond to survey or test questions. Although there will never be a perfect way to quantify a hidden variable, but one's traits may be obtained if questions on a single underlying dimension are asked (Bond & Fox, 2015), and tools such as CTT and IRT can be used in this regard. However, these measurement tools, though with some similarities, differ noticeably in their modelling procedures, characteristics of a measured construct and how respondents will react to an item (Sharkness & DeAngelo, 2011).

CTT predicts the expected results of difficulty of the items or the performance of test takers. The observed score (X) is the total sum of a true score (T) and an error score (E). Thus, standard test theory assumes that the observed score equals the sum of the correct score and errors: X = T + E. In contrast, IRT is a theory that uses mathematical models to assess talents, attitudes, and other characteristics through surveys and tests. According to Bond and Fox (2015), the likelihood of correctly identifying an item is dependent on the person and item factors. Trait level and item difficulty are inextricably linked; a tough item necessitates a higher trait level for a favorable response, whereas an easy item merely requires a low trait level to be replied to more quickly. This theory considers both the difficulty of the items and the performance of the examinees, offering sample-free statistics on both. Advantages include the ability to translate ordinal data into an interval scale, as well as the invariance property, which yields standard error of measurement (SEM) for each person or threshold/item parameter.

Rasch Model

Rasch Model

Georg Rasch, a Danish Mathematician, proposed the Rasch Model in 1960 among the IRT models. According to the model, a person's ability (Θ) and the difficulty of an item determine whether they correctly reply to or endorse it. Rasch recognized that defining the difficulty of an item and individual aptitude independent of which things individuals could answer were critical to developing objective measures in psychology. This model was then adopted by a wider collection of measurement models known as IRT, which has been widely used in educational research (Khine 2020). Researchers have utilized this to investigate psychometric qualities and address response bias (Ahn et al., 2023; Stolt, 2022). Rasch's use of the logit ruler satisfies Mok and Wright's five principles for human science, which include providing a linear measure, resolving missing data, offering precision estimates, analyzing outliers or misfits, and replicability (Sumintono 2018). When applied to attitudes, higher scores imply greater agreement among respondents and item

difficulty indicates how simple it is to endorse (Andrich 2013).

The Rasch Rating Scale Model, established by Andrich (1978), is excellent for Likert-scale because it ties a person's latent attribute to the likelihood of their response on a single scale. In this case, those with more of the latent feature are more likely to agree with an item than those with less of it. Through analyzing their position in the person/item map, meaningful inference can be made about both person and item interactions. Rasch models allow for assessment of the quality and usefulness of each rating scale item as well as providing an item hierarchy for the construct (Bond & Fox, 2015). The model assumes that affirmation of an item is based on the relative space between a person's level (e.g., job satisfaction) and item difficulty – where higher logit items require higher levels (Boone et al., 2014). Item parameters also express how difficult an item is to agree with, meaning that someone needs to have a high level of job satisfaction to strongly agree with an intense statement (Bond & Fox 2015). Each parameter is assigned a Standard Error Measurement (SEM).

Standard Error Measurement (SEM)

Birbaum (1968) defined SEM as an indicator of a test or scale item's correctness and trait level under consideration. This metric is not affected by the number of examinees and it provides insight into how much error is associated with ability estimates for different points on the ability continuum. Each item contributes its own amount to this measurement, independent from any other items on the test, which is different from traditional reliability estimates where an item's contribution to the overall test validity varies based on other items included in it. With Classical Test Theory (CTT), one item's influence on overall test reliability can be hard to detect. In the context of Rasch's Rating Scale Model, each item has its own difficulty index or threshold and standard error of measurement (SEM) associated with it. Generally, an SEM value under 0.5 is seen as low (i.e., high in reliability) while higher values are seen as high (low in reliability). Therefore, a lower SEM typically denotes better estimation ability by that item (Baumgartner, 2012). Responses given by persons to individual items can help distinguish high and low difficult items as well.

Person & Item Separation Indices

Person & Item Separation Indices are used to discover how people who responded items in an instrument were able to organize them in a hierarchy, from more difficult to easier level of acceptance, and if the instrument items were able to differentiate between people (teachers) with higher or lower levels of a trait (Bond & Fox, 2015). The Person Separation Index (PSI) examines whether the scale can manage to distinguish between different levels of job satisfaction. Low person separation in a given sample indicates that the instrument may lack the sensitivity to distinguish between high and low performers; additional questions may be required. The item separation index is used to assess the item hierarchy, specifically if each item contributes uniquely to comprehending the variable. Low item separation indicates that there are insufficient persons or larger samples are required to establish the instrument's item

difficulty hierarchy (construct validity). Item response theory models have some assumptions such as unidimensionality.

Unidimensionality

The Rasch Model, like other IRT models, is built on the premise of unidimensionality, suggesting that items on a test measure the same dimension. This is congruent with the concept used in physical sciences which adopts measuring a construct at a time, as is done in physical sciences (Bond & Fox, 2015). Furthermore, Rasch models are based on the idea that an underlying latent characteristic influences test outcomes (Boone, Staver, & Yale, 2014). This latent feature can be found in various levels among a target group (Ojerinde, 2013) and are frequently used to indicate that a single dominating talent or trait underpins performance on several tests or scales.

Method

For this study, a combination of equipment and survey research designs were used. Enugu State, one of 36 states in Nigeria was chosen by simple random sampling (balloting). According to data released by the Post Primary School Management Board's Planning, Research, and Statistics (PRS) department for 2019, there were 7303 teachers in public secondary schools (2,568 males and 4,735 females). Following that, a multi-stage sampling technique was utilized to choose a sample of 972 secondary school teachers, including 375 males and 597 females.

An instrument was designed for this study will enable data collection. The "Teachers' Job Satisfaction Scale (TJSS)" is made up of only 90 items (on six variables), allowing individual instructors to express their sentiments and attitudes toward their professions in a precise and efficient manner. This instrument was made up of two components. Part I requested that teachers supply personal information such as their gender and location. Part II featured 90 items on various themes, each of which was graded on a 4-point Likert scale (from "strongly agree" to "strongly disagree") based on whether the statement was favorably or negatively cued.

Instrument which comprised 90 items on TJSS was initially administered to 50 teachers different from the study sample, and all copies were later recovered. Individual item responses were subjected to factor analysis to confirm that the items were valid, adequate, and accurate representations of the construct. Using the Rotated Component Matrix, items were loaded on four renamed factors: interpersonal interaction (9 items), work itself (15 items), working condition (21 items), and remuneration (10 things). Using Schuster and Milland's (1978) minimum factor loading threshold of.350, 23 items were judged to be factorially impure while 12 items were found to be factorially complex. Thus, 35 items were removed following factor validation, leaving 55 for TJSS at this point. A Cronbach alpha test was then performed for the entire instrument. Interpersonal interaction (.70), Work itself (.77), Working condition (.81), and Remuneration (.72) all had coefficient alphas of.72. The total instrument's coefficient alpha was .87, indicating consistency which is in line with Cronbach and Shavelson's (2004) statement that intercorrelations across test items are highest when they all measure the same construct.

AJEMATES
African Journal of Educational Management, Teaching and Entrepreneurship Studies
VOL.18 (1), May-August, 2026; https://ajemates.org

The 55-item measure was distributed to a sample of 972 instructors drawn from 36 schools. They coordinated with the principals of those schools to make the questionnaire distribution and collation easier. The data was analyzed using WINSTEPS 5.1.1 software and the Rasch Rating Scale Model. The Conditional Maximum Likelihood Estimation technique was used to estimate item parameters, and it also generates scores for each item, as well as their standard errors. Using Winsteps measurement software, a Principal Components Analysis (PCA) of standardized residual correlations was conducted to evaluate the dimensionality. According to Adedoyin (2010), if a single underlying factor is present, unidimensionality is suggested.

Results and Discussion

Table 1: Table of Standardized Residual Variance (in Eigenvalue units)

		-- Empirical --		Modeled
Total raw variance in observations	=	89.5	100.0%	100.0%
Raw variance explained by measures	=	**44.5**	**48.7%**	**49.5%**
Raw variance explained by persons	=	20.7	22.3%	22.6%
Raw Variance explained by items	=	23.8	26.4%	26.9%
Raw unexplained variance (total)	=	45.0	51.3%	100.0% 50.5%
Unexplained variance in 1st contrast	=	**1.9**	**2.3%**	6.5%

Result from Table 1 showed that 48.7% of variance was explained by the Rasch dimension; 22.3% by person measures and 26.4% by item measures, with 51.3% left unexplained. To examine unidimensionality, Linacre (2014) stated that eigenvalues above 2 suggest that the extracted dimension is a secondary one which would threaten unidimensionality, however in this study the first contrast only had an eigenvalue of 1.9 signifying that it was unidimensional; thus, all items of the TJSS measure a single construct.

Table 2: Rasch Fitting Model; Infit and Outfit Mean Square (MNSQ) of the items

ITEM NUMBER	SAMPLE SIZE(n)	INFIT MNSQ	OUTFIT MNSQ
1	972	.75	.75
2	972	1.05	1.07
3	972	.85	.88
4	972	1.13	1.14
5	972	.91	.95
6	972	.85	.91
7	972	.78	.79
8	972	.75	.79
9	972	1.01	1.02
10	972	1.00	1.01
11	972	1.02	1.00
12	972	.83	.83

13	**972**	**2.14**	**2.13**
14	972	1.13	1.14
15	972	1.08	1.08
16	972	1.37	1.37
17	972	.94	.95
18	972	.93	.93
19	972	1.12	1.11
20	972	1.07	1.08
21	972	.92	.92
22	972	1.20	1.19
23	972	.72	.73
24	972	.87	.90
25	972	.99	1.02
26	972	.87	.93
27	972	.88	.89
28	972	1.28	1.32
29	972	.75	.76
30	**972**	**1.66**	**1.65**
31	972	1.10	1.09
32	972	1.02	1.03
33	**972**	**1.78**	**1.79**
34	972	.85	.84
35	972	.62	.62
36	972	.67	.69
37	972	1.06	1.10
38	972	.85	.85
39	972	.81	.84
40	972	1.05	.99
41	972	.92	.90
42	972	.90	.88
43	972	.98	.95
44	972	1.38	1.36
45	972	.96	.96
46	972	1.15	1.14
47	972	1.07	1.07
48	972	1.13	1.12
49	972	1.38	1.38
50	972	1.05	1.09
51	972	1.34	1.27
52	972	1.07	1.07
53	972	1.13	1.17
54	972	1.15	1.16
55	972	1.21	1.23
3. Overall Infit and Outfit Mean		**0.99**	**1.00**

AJEMATES
African Journal of Educational Management, Teaching and Entrepreneurship Studies
VOL.18 (1), May-August, 2026; https://ajemates.org

The results of the fit analysis for the TJSS items are presented in Table 2. The findings indicated that nearly all the items analyzed exhibited mean square fit statistic values falling within the range of .6 to 1.4, as established by Bond et al. (2020). A total of fifty-two (52) items examined, representing 95%, exhibited mean square infit/outfit values ranging from .62 to 1.38. In contrast, the remaining three items (5%) displayed mean square values exceeding 1.40, suggesting they were under-fitting. Items 13, 30, and 33, exhibiting infit and outfit mean square values of 2.14, 1.66, and 1.78; and 2.13, 1.66, and 1.79 respectively, were identified as misfitting and subsequently removed from consideration. The average of all infit mean square values was .99, while the average of all outfit mean square values was 1.00.

Table 3: Threshold Parameter Estimates of the Items and Model Standard Errors

Item number	Item Measure (Difficulty Index)	Model Standard Errors
51	1.73	.05
54	1.49	.04
48	1.47	.04
19	1.45	.04
26	1.41	.04
55	1.30	.04
50	1.23	.04
22	1.05	.04
13	1.03	.04
14	.89	.04
53	.89	.04
46	.87	.04
47	.84	.04
45	.81	.04
37	.71	.04
10	.60	.04
16	.44	.04
15	.38	.04
12	.36	.04
2	.36	.04
52	.29	.04
4	.24	.04
32	.21	.04
20	.18	.04
17	.10	.04
29	.07	.04
23	.05	.04
24	.04	.04
3	-.08	.04
49	-.010	.04
27	-.24	.04
21	-.30	.04
39	-.33	.04

AJEMATES
African Journal of Educational Management, Teaching and Entrepreneurship Studies
VOL.18 (1), May-August, 2026; https://ajemates.org

35	-.38	.04
38	-.43	.04
25	-.44	.04
33	-.44	.04
18	-.45	.04
28	-.59	.04
36	-.60	.04
31	-.71	.04
6	-.74	.04
30	-.74	.04
43	-.93	.05
34	-.96	.05
44	-.99	.05
5	-1.01	.05
7	-1.04	.05
41	-1.08	.05
9	-1.12	.05
40	-1.15	.05
42	-1.22	.05
8	-1.28	.05
11	-1.54	.05
1	-1.59	.05
Overall Threshold and Standard Error Mean:	**.00**	**.04**

Table 3 indicated that the difficulty levels of the items varied from -1.59 for item 1 to 1.73 for item 51, with standard errors ranging from .04 to .05. Approximately 50.9% of the items presented a moderate level of difficulty in endorsement (positive measures), whereas 49.1% were characterized by ease of endorsement (negative measures). The mean item measure, or overall threshold value, was 0, indicating that there were no excessively difficult items, as all values were 3 or below (Bond & Fox, 2015). The findings align with those of Natsir (2022), indicating that the item difficulty spans from -.71 to 3.05. This, however, does not align with the findings of the study by Marais, Styles & Andrich (2011), which reported no items with negative values and threshold parameters that ranged from 1.93 to 3.34, all of which were highly positive. The items in this study are precisely designed to assess the level of job satisfaction among secondary school teachers.

Furthermore, each item exhibited standard errors below 0.5, indicating a strong level of reliability in the measurements of the individual (Linacre, 2014). The SEMs for the item measures were notably small and consistent, ranging from .04 to .05, with an average of .04, indicating a high level of measurement accuracy for these items. This indicates that the items demonstrate a high level of reliability. The results align with Xavier's (2012) study, which reported a mean SEM for items of .06.

AJEMATES
African Journal of Educational Management, Teaching and Entrepreneurship Studies
VOL.18 (1), May-August, 2026; https://ajemates.org

Table 4: Person and Item Separation Indices

Category	Real Separation	Model Separation	Reliability
Person	2.16	2.39	0.85
Item	10.08	10.75	0.89

Table 4 illustrates the separation measures for both 'Real' and 'Model'. The Real can be viewed as representing the most unfavorable estimates, whereas the Model reflects the most favorable estimates, with the true distinction residing in the intermediate range (Edkins & Royal, 2011). The estimates for separation among individuals in the sample varied between 2.16 and 2.39, accompanied by a reliability index of 0.85. The Model Separation estimate for individuals in the sample was 2.39, exceeding the criterion of >2 (Linacre, 2021) and demonstrating sufficient spread; this is comparable to the findings of Zachariah et al. (2022) and Xavier (2012), whose person separation indices were 4.67 and 10.86, respectively. However, this finding contrasts with the research conducted by Marais, Styles, and Andrich (2011), which reported a PSI of less than 1, specifically 0.91. This suggests that individuals were not categorized into high and low levels of the trait in their study.

Furthermore, the Rasch item separation index varied between 10.08 and 10.75, accompanied by a reliability index of 0.89, exceeding the established criterion of >3 (Linacre, 2021). This suggests that the items of the instrument are either well-aligned or organized hierarchically, ranging from the most challenging to endorse to the least challenging. This observation aligns with Xavier's (2012) item separation indices of 3.88, yet it contradicts the findings of Zachariah et al. (2022), who reported an index of 2.42 for this specific measure.

Match Between Persons and Items

Figure 1 provides a visual representation of humans and objects along a continuum, enhancing the understanding of the distinction between individuals and the arrangement of items.:-

AJEMATES
African Journal of Educational Management, Teaching and Entrepreneurship Studies
VOL.18 (1), May-August, 2026; https://ajemates.org

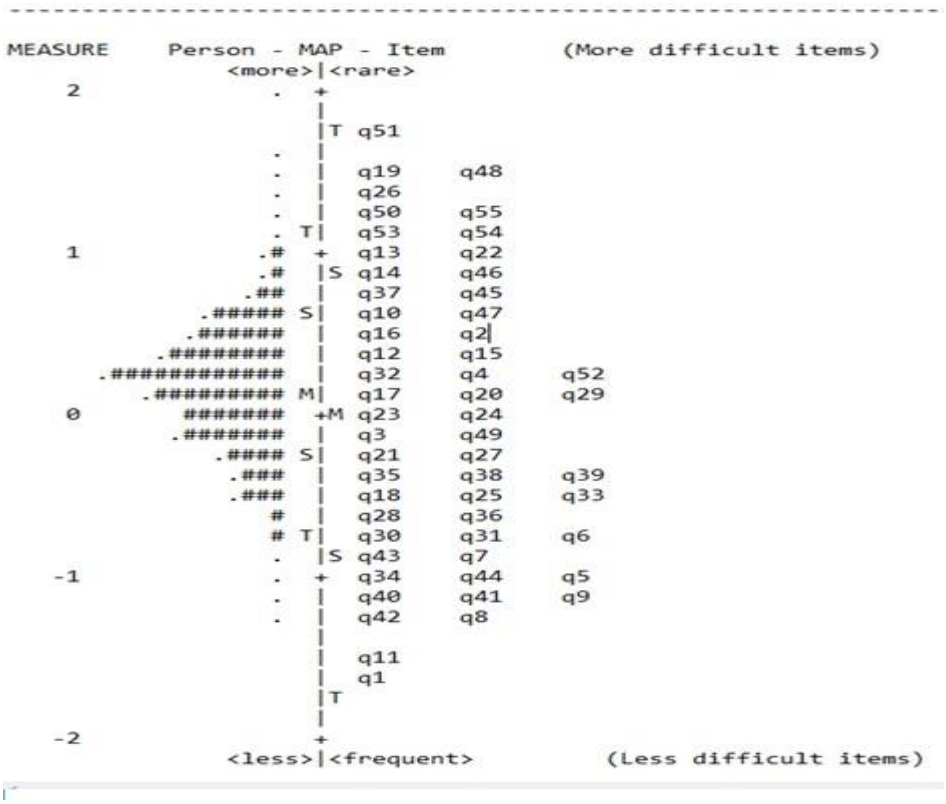


Fig 1: Person/Item Map

Figure 1 illustrates a map categorizing the items according to their endorsement difficulty, alongside the respondents' corresponding willingness to endorse them. To the left of the item map, there exists a logit scale that ranges from +2 to -2 in a descending order. The map is divided into two sections, featuring individuals on the left and survey items on the right, both arranged along the same logit scale in either ascending or descending order; person respondents are denoted with "# (n=12)" or "." (n=1 to 11). Advancing on the map increases the challenge of endorsing items; findings indicated that the most straightforward items were "With teaching I can help my students learn" (Q11) and "I have a good relationship with my students" (Q1), whereas "My salary is less than what I deserve" (Q51) proved to be the most challenging. Bond and Fox (2015) observed that the absence of gaps between items allowed for an adequate measurement of respondents' ability or trait levels, suggesting that the TJSS effectively measures varying levels of job satisfaction.

Conclusion

Based on the outcomes of this study, it was established that the TJSS is a valid measure of teacher job satisfaction that can account for a wide range of individual ability and item difficulty. The item hierarchy was well-structured from higher to lower difficulty, and teachers were appropriately classified as having high or low levels of job satisfaction. This suggests that the TJSS can be used to assess and

document teachers' job satisfaction when performing their jobs. Therefore, the TJSS could prove to be a useful tool for the Ministry of Education to identify any deficiencies in policies so as to plan newer professional development strategies for improving teaching skills. Psychometricians were advised to use IRT methods to evaluate the psychometric characteristics of a scale on an item-by-item basis rather than relying on a single index, as is commonly done in CTT.

Declarations

Ethics Approval and Consent to Participate

The authors obtained the approval from the Post Primary School Management Board, Enugu state, Nigeria.

Consent for Publication

The authors have read and consent for the final version of this manuscript to be published.

Availability of data and materials

All the data used in the study are available upon request.

Conflicts of interest statement

The authors declare no conflict of interest, financial or otherwise.

Funding

The authors received no funding for this study.

Author Contributions

Metu, I,C. – main idea, conceptualization, data collection, data analysis and writing of the manuscript.

Madu, B.C. – conceptualization, review and validation.

Metu C.A. – data collection, reporting and editing.

REFERENCES

Ahn, S., Pinheiro, P. S., McClure, L. A., Hernandez, D. R., Caban-Martinez, A. J., & Lee, D. J. (2023). An examination of psychometric properties of study quality assessment scales in meta-analysis: Rasch measurement model applied to the firefighter cancer literature. *PloS one*, *18*(7), e0284469. https://doi.org/10.1371/journal.pone.0284469

Andrich, D. (1978). Application of a psychometric rating model to ordered categories which are scored with successive integers. *Journal of Applied Psychology*, *2*(4), 581-594. https://doi.org/10.1177/014662167800200413

Andrich, D. (2013). *An expanded threshold structure of the polytomous rasch model that dispels any "Threshold disorder controversy*" San Diego, CA: Academic Press.

Baumgartner T.A., (2012). *Conducting and reading research in health and human performance.* 3rd Edition. NewYork: MC-Gram Hill High Education.

Bichi, A. A., &Talib, R. (2018). Item Response Theory: An Introduction to Latent Trait Models to Test and

Item Development. *International Journal of Evaluation and Research in Education*, *7*(2), 142-151. https://eric.ed.gov/?id=EJ1183755

Birnbaum, A., (1968). Some latent trait models and their use in inferring an examinee's ability, In FM Lord & M.R Norvick (Eds.). *Statistical Theories of Mental Test Scores.* Reading MH: Addison-Wesley.

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences*. New York, NY: Routledge.

Bond, T. G., Yan, Z., & Heene, M. (2020). *Applying the Rasch model: Fundamental measurement in the human sciences*. (4th ed.). Routledge.

Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch Analysis in the human sciences*. Dordrecht, Springer. http://dx.doi.org/10.1007/978-94-007-6857-4

Chaudhary, V., Vemuri, V. P., Cavaliere, L. P. L., Verma, V., Manoharan, G., & Bharti, A. (2024, March). A comparative analysis of job satisfaction and motivational factors of employees in public versus private organisations. *In AIP Conference Proceedings 2816*(1). AIP Publishing. https://doi.org/10.1063/5.0186575

Cheng, K., Cao, X., Guo, L., & Xia, Q. (2022). Work connectivity behavior after-hours and job satisfaction: examining the moderating effects of psychological entitlement and perceived organizational support. *Personnel Review*, *51*(9), 2277-2293. http://dx.doi.org/10.1108/PR-06-2020-0413

Edkins, V., & Royal, K. (2011). Psychometric properties of death anxiety scale (DAS) among terminally ill cancer patients. *Journal of Psychosocial Oncology*, 29(4), 359-371. https://www.tandfonline.com/doi/abs/10.1080/07347332.2011.582639

Federal Republic of Nigeria (2013). *Teachers' Registration Council of Nigeria Handbook.* Nigeria.

Gall, M.D., Borg, W.R., & Gall, J.P. (2015). *Educational research: An introduction* (8th ed.). Longman.

Hand, B.N.; Velozo, C.A.; & Krause, J.S. (2017) Rasch measurement properties of the Pain Medication Questionnaire in persons with spinal cord injury. *Spinal Cord*, *55*, 1117–1122. https://www.nature.com/articles/sc201789

Khine, M.S. (2020). Objective Measurement in Psychometric Analysis. In: Khine, M. (eds) *Rasch Measurement*. Singapore. https://doi.org/10.1007/978-981-15-1800-3_1

Linacre, J.M. (2014). *Estimation methods for rasch measures*. In E.V. Smith & R. M. Smith (Eds.) *Introduction to rasch measurement*. JAM Press.

Linacre, J.M. Winsteps Rasch Measurement Computer Program: User's Guide (2021). Available online: https://www.winsteps.com/a/Winsteps-Manual.pdf

Locke, E. (1976). The nature and causes of job satisfaction. In M. D. Dunnette (Eds.), *Handbook of Industrial Organizational Psychology* (pp. 1297-1349). Rand-McNally.

Marais, I., Styles, I., & Andrich, D. (2011). Report on rasch analysis of the academic: Perceptions of the teaching environment (APOTE) for 2010. The University of Western Australia. https://bit.ly/3Fx53FM

Natsir, I., Mayasari, D., & Munfarikhatin, A. (2022). Rasch Model: Quality of test instruments to measure students' ability in algebra. *SHS Web of Conferences, 149*. EDP Sciences. https://doi.org/10.1051/shsconf/202214901033

Ojerinde, D. (2013). "Classical Test Theory (CTT) VS Item Response Theory (IRT): An Evaluation of the Comparability of Item Analysis Results". A guest lecture presented at the Institute of Education, University of Ibadan on 23rd May, 2013.

Pepe, A., Addimando, L., & Veronese, G. (2017). Measuring teacher job satisfaction: Assessing invariance in the teacher job satisfaction scale (TJSS) across six countries. *Europe's journal of psychology*, *13*(3), 396. https://psycnet.apa.org/doi/10.5964/ejop.v13i3.1389

Rasch, G. (1980). Probabilistic models for intelligence and attainment tests (expanded edition). University of Chicago Press (original work published in 1960).

Sharkness, J., & DeAngelo, L. (2011). Measuring students' involvement; a theoretical comparison of CTT and IRT. https://eric.ed.gov/?id=EJ930336

Stolt, M., Kottorp, A., & Suhonen, R. (2022). The use and quality of reporting of Rasch analysis in nursing research: A methodological scoping review. *International journal of nursing studies, 132*, 104244. https://doi.org/10.1016/j.ijnurstu.2022.104244

Sumintono, B. (2018). Rasch Model measurements as tools in assessment for learning. Proceedings of 1st International Conference on Education Innovation (ICEI 2017). Atlantis Press.

Tabatabaee-Yazdi, M., Motallebzadeh, K., Ashraf, H., & Baghaei, P. (2018). Development and Validation of a Teacher Success Questionnaire Using the Rasch Model. International Journal of Instruction, 11(2), 129-144. https://doi.org/10.12973/iji.2018.11210a

Tamsen, C.V., (2017). Using rasch analysis to inform rating scale development. Res High Educ (pp. 50-55). Pergamon Press.

Wright, B.D. Unidimensionality coefficient. Rasch Meas. Trans. 1994, 8, 385.

Wright, B. & Mok, M. (2004). *An overview of family of Rasch measurement models*. Jam Press.

Xavier, W. J., (2012). *Measuring job satisfaction among Kentucky head*

principals using the rasch rating scale model. [Unpublished Ph.D Thesis], University of Kentucky.

Zechariah, S., Waller, J. L., Stallings, J., Gess, A. J., & Lehman, L. (2022). Item-level analysis of a newly developed interactive nutrition specific physical exam competency tool (INSPECT) using the rasch measurement model. *Healthcare 10*(2), 259-259. https://doi.org/10.3390/healthcare10020259

Zhou, X., Padrón, Y., Waxman, H. C., Baek, E., & Acosta, S. (2024). How do School climate and professional development in multicultural education impact job satisfaction and teaching efficacy for STEM teachers of English learners? A path-analysis. International Journal of Science and Mathematics Education, 22(2), 447-468. https://doi.org/10.1007/s10763-023-10381-y