Determining Dimensionality and Model-Data Fit Statistics of the ICT Test Using the Rasch Model at Jigawa State University

by

Abubakar Rabiu Uba[1]

Ahmad Zamri Khairani[2]

[1]Department of Education, Sule Lamido University Kafun Hausa

abubakar.rabiu@slu.edu.ng, 08061679685

[2]School of Educational Studies, Universiti Sains Malaysia

ahmadzamri@usm.my, +60195492525.

Abstract

This study evaluates the dimensionality and model-data fit statistics of an Information and Communication Technology (ICT) test administered at Jigawa state university using the Rasch model. Despite the test's widespread use, its psychometric properties had not been previously assessed. This cross-sectional study addresses this gap by employing Rasch analysis to evaluate the reliability and validity of the test items, using a sample of 600 undergraduate test scripts randomly selected from various departments. This approach ensures that the measurement instrument accurately reflects the underlying construct of ICT proficiency across a diverse student population. The results of the Principal Component Analysis indicate that the variance explained by the measures was 62.9%, confirming a strong primary dimension and supporting the one- dimension nature of the test. Additionally, model-data fit statistics shows that the mean values for both infit and outfit mean-square (MNSQ) were 1.00 and 1.03 logits, respectively, aligning with the expected values for the Rasch model. These findings confirm that the test items behave as expected, demonstrating strong psychometric properties in terms of validity and reliability. The findings reveal that the test is unidimensional and possesses strong fit statistics, confirming its validity and reliability. These results provide valuable insights into the test's psychometric properties, validating its effectiveness in measuring ICT proficiency across different academic disciplines. Consequently, the study supports the continued use of the test while also offering guidance for potential improvements in its design and application within educational settings.

Keywords: Dimensionality, fit statistics, rasch analyses, ict test.

Determining Dimensionality and Model-Data Fit Statistics of the ICT Test Using the

Rasch Model at Jigawa State University

In the rapidly evolving landscape of higher education, Information and Communication Technology (ICT) proficiency has transitioned from a desirable skill to an essential competency for students. This shift is particularly pronounced in developing countries like Nigeria, where mastery of digital technologies is a catalyst for economic growth, innovation, and global competitiveness (Olaniyi & Ismaila, 2016). The integration of ICT into educational curricula represents more than just a modernization of teaching methods; it embodies a fundamental transformation in how knowledge is accessed, processed, and applied in both academic and professional contexts.

The imperative for ICT literacy in Nigerian universities is underscored by the country's aspirations to bridge the digital divide and position itself as a key player in the global knowledge economy. As noted by Adebisi et al. (2018), the Nigerian government has increasingly recognized the pivotal role of ICT in driving sustainable development, and has consequently prioritized digital literacy in its national education policy. This emphasis has cascaded down to the tertiary level, where universities are tasked with not only imparting subject-specific knowledge but also equipping students with the technological skills demanded by an increasingly digitized workforce.

However, the rapid adoption of ICT in Nigerian higher education institutions has outpaced the development of robust assessment frameworks. While various tools have been implemented to measure students' technological competencies, the psychometric properties of these instruments often remain unexamined or inadequately validated (Adedokun-Shittu & Shittu, 2015). This gap in knowledge raises critical questions about the reliability and validity of ICT assessments across diverse student populations, particularly in a country as ethnically and linguistically diverse as Nigeria. The implications of this oversight are far-reaching. In an educational context where ICT proficiency can significantly influence academic outcomes and future career prospects, the use of invalidated assessment tools may lead to in-accurate evaluations of students' capabilities. This, in turn, could result in mis-aligned educational interventions, in-equitable academic progression, and

ultimately, a mis-match between graduate skills and industry requirements (Olelewe & Agomuo, 2016). At the heart of this issue lies the GSP122 ICT test administered at a public university in Nigeria, which serves as the focus of this study. Despite its widespread application across various departments at Jigawa State University and other public universities in Nigeria, the test's psychometric qualities in terms of dimensionality and model-data fit statistics have not been previously scrutinized. This lack of psychometric evaluation represents a critical blind spot in the university's assessment strategy. As emphasized by Messick (1995), the validity of an educational assessment extends beyond its content to encompass its internal structure, response processes, and consequences of use. Therefore, a comprehensive examination of this ICT test's psychometric properties is not merely an academic exercise but a fundamental requirement for ensuring its efficacy and fairness.

The Rasch model, a sophisticated psychometric approach within Item Response Theory, offers a powerful framework for addressing these concerns. Developed by Georg Rasch (1960), this model provides several advantages over classical test theory approaches. Firstly, it allows for the conversion of ordinal data into interval measures, enabling more precise and meaningful comparisons of both item difficulties and person abilities. Secondly, the Rasch model's invariance property ensures that item parameters remain consistent across different subgroups of respondents, a crucial feature when assessing diverse student populations (Bond and Fox, 2015).

Moreover, the Rasch model's one- dimension nature's assumption provides a rigorous test of whether the ICT assessment measures a single, coherent construct. This is particularly relevant in the context of ICT proficiency, which encompasses a range of skills from basic computer literacy to advanced problem-solving using digital tools. By examining the dimensionality of the test, we can determine whether it effectively captures the complexity of ICT competence or if it inadvertently conflates distinct skill sets.

The present study aims to address these critical gaps by employing Rasch analysis to evaluate the dimensionality and model-data fit statistics of the GSP122 (ICT test) for the 2023/2024 academic session at Jigawa state university. By examining a sample of 600 undergraduate test scripts randomly selected from various departments, this study seeks to:

1. Assess the one- dimension nature of the ICT test, ensuring that it measures a single, coherent construct of ICT proficiency. This analysis will reveal whether the test items collectively contribute to measuring the same underlying trait or if there are distinct sub-domains within the assessment.

2. Evaluate the model-data fit statistics to determine how well individual test items align with the Rasch model's expectations. This process will identify any problematic items that may be measuring constructs other than ICT proficiency or that function differently across sub-groups of students.

The findings of this study have profound implications for both educational practice and policy in Nigeria and potentially other developing countries facing similar challenges in ICT education. A rigorously validated ICT assessment tool can serve multiple purposes: Firstly, it can inform curriculum development by highlighting areas where students consistently struggle, enabling educators to tailor their instructional strategies to address specific weaknesses in ICT proficiency. Secondly, it can guide the allocation of educational resources, ensuring that investments in ICT infrastructure and training programs are targeted where they are most needed. Thirdly, a validated assessment can enhance the credibility of ICT certifications awarded by the university, potentially improving graduates' employability in an increasingly competitive job market.

Moreover, in the broader context of Nigerian higher education, this research contributes to the on-going efforts to enhance the quality and relevance of ICT education in preparing students for the digital workforce. As noted by Olelewe and Agomuo (2016), there is a growing concern about the mis-match between the ICT skills taught in universities and those required by employers. By ensuring that ICT assessments accurately measure relevant competencies, universities can better align their curricula with industry needs, thereby enhancing the employability of their graduates.

From a methodological perspective, this study demonstrates the application of advanced psychometric techniques in evaluating educational assessments in a developing country context.

The use of Rasch analysis not only provides a more nuanced understanding of the ICT test's properties but also sets a precedent for rigorous evaluation of other educational assessments in Nigeria. This approach aligns with global trends towards evidence-based educational practices and can contribute to raising the overall standard of assessment in Nigerian higher education.

By rigorously examining the psychometric properties of this test, this study not only validates its continued use but also provides a methodological framework for evaluating similar assessment tools in other educational contexts. The results offer valuable insights that can guide potential improvements in the test's design and application, ultimately enhancing the measurement of ICT proficiency in Nigerian universities and beyond. Therefore, as Nigeria and other developing nations strive to harness the power of ICT for national development, the importance of accurate and fair assessment of digital competencies cannot be overstated. This study represents a crucial step towards ensuring that the tools used to measure ICT proficiency are subject to rigorous scientific scrutiny. The insights gained from this research have the potential to inform policy decisions, improve educational practices, and ultimately contribute to the development of a more digitally literate and globally competitive workforce in Nigeria.

Methods

This study employed a cross-sectional research design to evaluate the dimensionality and model-data fit statistics of an Information and Communication Technology (ICT) test administered at Jigawa state university. The cross-sectional approach was chosen as it allows for the collection of data from a diverse sample of participants at a single point in time, providing a snapshot of the test's psychometric properties across various academic disciplines (Levin, 2006).

The study sample consisted of 600 undergraduate students from Jigawa State University. Participants were randomly selected from various departments to ensure a representative sample. The random sampling technique was implemented to minimize selection bias and enhance the generalizability of the findings (Etikan & Bala, 2017). The sample size of 600 was determined based on recommendations for Rasch analysis, which suggest a minimum of 250-500 respondents

for stable item calibrations (Linacre, 1994). The large sample size of 600 was chosen to account for potentially unusable data and, to ensure adequate representation across disciplines.

The instrument under assessment was the university's standardized ICT proficiency test. This test is routinely administered to undergraduate students as part of their academic assessment. The test consists of 60 multiple-choice items designed to measure various aspects of ICT proficiency, including basic computer skills, software applications, internet usage, etc. Data collection was conducted during the 2023/2024 academic period. The researchers obtained permission from the university administration to access the test scripts of the randomly selected participants. To ensure confidentiality and ethical compliance, all personal identifiers were removed from the test scripts before data analysis, and each script was assigned a unique code. The ICT test was administered under standardized conditions as part of the university's regular assessment process. Students completed the test in computer laboratories and other lecture rooms under supervision, with a time limit of 1 hour and 30 minutes.

The data analysis was conducted using Winsteps version 4.8.0, a software package specifically designed for Rasch measurement analysis (Linacre, 2021). The analysis proceeded in two stages: Firstly, Dimensionality Assessment: Principal Component Analysis (PCA) of residuals was conducted to evaluate the one- dimension nature of the ICT test. The criteria for one-dimension nature were set as the percentage of the raw variance explained by the measures should be at least 20% (Reckase, 1979) or 40% as suggested by Linacre (2006), the eigenvalue should be less than 3.0 for the first contrast and less than 10% of the variance explained by the first contrast (Linacre, 2021). Secondly, Item Fit Statistics: Infit and outfit mean-square statistics were calculated to assess how well individual items fit the Rasch model. Items with infit and outfit mean-square values between 0.6 and 1.4 were considered to have an acceptable fit (Bond & Fox, 2015). 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 provides 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 one- dimension nature 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 3.0 for the first extracted construct is considered to violate the unidimensional assumption (Linacre, 2002). According to the same author, a scale demonstrates one- dimension nature when the primary measurement dimension accounts for greater than 40% of the total variance, indicating that the scale effectively measures a single underlying construct.

Furthermore, in our examination of the ICT proficiency test's psychometric properties, we employed Baghaei's (2008) framework to analyze evidence supporting the construct validity of this measure. This approach provides a comprehensive lens through which to evaluate the validity of assessment instruments, particularly in educational contexts. Construct validity, a fundamental concept in psychometrics, refers to the degree to which a test measures what it claims to measure. In the case of our ICT proficiency test, construct validity is crucial to ensure that the assessment accurately reflects students' true ICT competencies, rather than extraneous factors. Baghaei's framework emphasizes the importance of identifying and addressing threats to construct validity. One such threat, which our study focused on, is construct-irrelevant variance. This concept refers to the extent to which test scores are influenced by factors unrelated to the construct being

measured. In essence, construct-irrelevant variance introduces "noise" into the measurement, potentially leading to inaccurate assessments of students' ICT skills.

To investigate construct-irrelevant variance, we will utilize the previously discussed infit and outfit mean-square (MNSQ) statistics in our model-data fit analysis. These statistics, derived from Rasch analysis, provide valuable insights into how well individual test items align with the overall construct being measured. The infit MNSQ is particularly sensitive to unexpected response patterns for items that are closely matched to the respondent's ability level. In contrast, the outfit MNSQ is more sensitive to outliers – unexpected responses to items that are either much easier or more difficult than the respondent's ability level. Both statistics play a crucial role in identifying items that may be introducing construct-irrelevant variance. In this study's analysis, we considered items with infit and outfit MNSQ values between 0.4 and 1.6 to have acceptable fit, as suggested by Bond and Fox (2015). Items falling outside this range were flagged for further examination, as they might be measuring constructs other than ICT proficiency or functioning differently across subgroups of students.

Results

For the dimensionality, table 1 presents the results of the Principal Component Analysis (PCA) of the standardized residuals for the GSP122 ICT test. In Rasch analysis, the percentage of the raw variance explained by the measures is one key indicator of one- dimension nature; therefore, the PCA results show that the variance explained by the measures was 62.9%, indicating a strong primary dimension. The value exceeds the minimum acceptable threshold of at least 20% set by Reckase (1979) and 40% set by Linacre (2006), thus confirming that the GSP122 test items measure only one unidimensional construct, namely ICT knowledge or proficiency. Furthermore, Principal Component Analysis (PCA) reveals that the degree of interference on the items being measured or an unexplained variation in the first contrast is recorded at 5.4%, which is within the good range of 5 to 10% (Fisher, 2007). In addition, numbers below 10% indicate compliance on a one-dimensional scale, as stated by Linacre (2007) and Eakman (2012). On the other hand, the results show an eigenvalue of 3.2 logits, which suggests the existence of a minor secondary

dimension that does not necessarily call into question the GSP122 test's overall one-dimensionality (Linacre, 2002).

Table 1

Results from the principal component analysis of residuals

			Empirical		Modelled
Total raw variance in observations	=	161.7	100.0%		100.0%
Raw variance explained by measures	=	101.7	62.9%		63.0%
Raw variance explained by persons	=	81.9	50.7%		50.8%
Raw Variance explained by items	=	19.8	12.2%		12.2%
Raw unexplained variance (total)	=	60.0	37.1%	100.0%	37.0%
Unexplained variance in 1st contrast	=	3.2	5.4%	6.4%	
Unexplained variance in 2nd contrast	=	3.0	4.0%	5.3%	
Unexplained variance in 3rd contrast	=	2.4	2.5%	3.9%	
Unexplained variance in 4th contrast	=	1.9	1.1%	3.1%	
Unexplained variance in 5th contrast	=	1.8	1.1%	3.0%	

With regards to Model-Data fit Statistics, the means for both infit and outfit MNSQ were found to be 1.00 and 1.03 logits, which are the expected values for both fit statistics. Thus, the items behave as expected according to the Rasch model. As shown in Table 2, when individual items were examined, it was discovered that infit MNSQ values ranged from .96 to 1.04 logits, while outfit MNSQ values ranged from .95 to 1.37 logits. The values are within the acceptable range of 0.6 to 1.4 logits based on the guidelines by Bond and Fox (2015). Thus, it can be concluded that the measures showed slight variation from model expectations. There was evidence of consistency between 600 examinees' responses on the 60 GSP122 test items, further confirming that the items possess good fit between the model's expectations and the empirical data. Additionally, all items' infit and outfit MNSQ values indicated an acceptable difference between empirical data and the model's assumptions in terms of construct validity. As such, it can be concluded that the threat to construct irrelevant variance is minimized.

Regarding Table 2, the Entry column indicates the order in which items were entered into WINSTEPS. The Total Score indicates the sum of scores obtained by adding up all the participants scores. Since GSP122 is a multiple-choice score, it can be safely said that 277 participants got the answer correctly. The Total Count showed the number of participants who attempted the item. It

shows that all items (60) were answered by all participants (600). Meanwhile, the Measure statistics refer to the item difficulty statistics estimated by WINSTEPS. The higher the value of the item measures, the more difficult the item is for the students. For example, Item 22 (measure = 2.69 logits) is more difficult compared to Item 40 and 43 (measure = 2.64 and .82 logits, respectively). In fact, Item 22 is considered the most difficult item in the GSP122 test, whereas Item 36 (measure = -0.41 logits) is the easiest.

Table 2

Item Statistics for GSP 122 ICT Test Items

Entry Level	Total Score	Total count	Measure	Model SE	Infit MNSQ	Outfit MNSQ	PT-Measure Corr.
Item22	277	600	2.69	0.08	1.00	1.00	.16
Item40	309	600	2.64	0.08	.99	.98	.21
Item43	291	600	0.82	0.07	1.02	1.00	.18
Item16	315	600	0.81	0.07	1.04	1.30	.13
Item41	316	600	0.81	0.07	1.02	1.26	.14
Item42	316	600	0.81	0.07	1.01	1.01	.16
Item2	292	600	0.79	0.07	1.03	1.05	.16
Item13	294	600	0.79	0.07	.99	.95	.17
Item28	303	600	0.76	0.08	1.02	1.33	.19
Item24	305	600	0.76	0.08	1.02	1.06	.20
Item1	425	600	0.72	0.08	1.01	1.37	.24
Item9	276	600	-0.10	0.08	1.02	1.02	.12
Item6	279	600	-0.12	0.08	.97	.97	.25
Item18	279	600	-0.12	0.08	1.02	1.02	.12
Item21	280	600	-0.13	0.08	1.01	1.01	.14
Item17	282	600	-0.14	0.08	1.01	1.01	.15
Item7	284	600	-0.15	0.08	1.00	1.00	.18
Item8	285	600	-0.16	0.08	1.01	1.01	.14
Item50	287	600	-0.17	0.08	.99	.99	.20
Item29	288	600	-0.18	0.08	1.00	1.00	.17
Item4	289	600	-0.19	0.08	.98	1.00	.22
Item5	289	600	-0.19	0.08	.99	1.01	.18
Item20	289	600	-0.19	0.08	1.01	1.01	.15
Item38	290	600	-0.19	0.08	.99	.99	.19
Item46	290	600	-0.19	0.08	.98	.98	.22
Item12	291	600	-0.20	0.08	1.02	1.01	.13
Item10	292	600	-0.21	0.08	1.00	1.01	.16
Item25	292	600	-0.21	0.08	1.00	1.00	.18
Item26	294	600	-0.22	0.08	1.00	1.00	.17
Item45	294	600	-0.22	0.08	.99	.99	.19

Item							
Item15	296	600	-0.23	0.08	1.02	1.02	.12
Item19	297	600	-0.24	0.08	1.03	1.04	.08
Item35	297	600	-0.24	0.08	1.01	1.01	.15
Item27	298	600	-0.25	0.08	.98	.98	.21
Item54	298	600	-0.25	0.08	1.01	1.00	.16
Item23	299	600	-0.25	0.08	.98	.98	.19
Item51	300	600	-0.26	0.08	.99	.99	.22
Item32	301	600	-0.27	0.08	1.00	1.00	.17
Item47	301	600	-0.27	0.08	1.03	1.03	.10
Item55	301	600	-0.27	0.08	1.02	1.02	.12
Item11	302	600	-0.28	0.08	1.00	1.00	.17
Item30	302	600	-0.28	0.08	1.01	1.00	.16
Item31	303	600	-0.28	0.08	1.02	1.02	.12
Item33	303	600	-0.28	0.08	1.02	1.02	.12
Item44	303	600	-0.28	0.08	.98	.97	.23
Item56	305	600	-0.30	0.08	.99	.99	.19
Item34	306	600	-0.30	0.08	.98	.97	.24
Item3	307	600	-0.31	0.08	.99	.99	.21
Item49	307	600	-0.31	0.08	1.00	1.00	.17
Item39	309	600	-0.32	0.08	1.02	1.02	.13
Item57	309	600	-0.32	0.08	.99	.98	.21
Item52	310	600	-0.33	0.08	.99	1.00	.20
Item37	312	600	-0.34	0.08	1.01	1.01	.13
Item60	314	600	-0.36	0.08	.97	.99	.23
Item53	315	600	-0.36	0.08	1.01	1.01	.15
Item59	315	600	-0.36	0.08	.99	.98	.21
Item14	317	600	-0.38	0.08	1.00	1.00	.17
Item48	320	600	-0.40	0.08	.98	.98	.21
Item58	320	600	-0.40	0.08	.96	.96	.27
Item36	321	600	-0.41	0.08	.99	.99	.19
Mean	303.3	600.0	0.00	0.08	1.00	1.03	
SD	19.1	.2	.62	0.01	.02	.12	

Discussion

The GSP122 test is a general examination administered at Jigawa state university, consisting of 60 multiple-choice questions. This part discusses the results for the following important parts of the test's psychometric properties. These are (a) model-data fit statistics, which check how well the data fits the Rasch model; and (b) dimensionality, which checks to see if the test measures a single construct or multiple constructs. Together, these analyses provide an impressive evaluation of the GSP122 test's psychometric qualities and its effectiveness as an assessment tool.

Dimensionality

Findings from this study provide strong evidence for the one- dimension nature of the GSP122 ICT test, confirming that it measures only one construct, namely ICT knowledge, as evidenced by the fit statistics and the variance explained by the measures. These findings could be explained by the students' perception that the items were understandable and connected to particular ICT knowledge tasks, which lessened the impact of guessing and cheating. This finding is consistent with recent studies that have employed Rasch analysis to examine the dimensionality of tests in various fields (Hamad, 2021; Paek & Wilson, 2018; Yang et al., 2020). The one-dimension nature of the GSP122 test supports the notion that the test measures a single underlying construct, which is a fundamental assumption of the Rasch model (Wright & Masters, 2018). This finding is also in line with the theoretical framework that underpins the test's development, which posits that ICT knowledge is a distinct construct (Smith, 2017).

The results of this study make significant contributions to the ongoing discourse on the importance of dimensionality in test validation. Bond and Fox (2015) emphasize that a test's dimensionality can substantially influence both the interpretation of test scores and the validity of inferences drawn from these scores. The finding of one- dimension nature in the GSP122 ICT test underscores the critical need to examine the dimensionality of assessment instruments to ensure the validity and interpretability of test scores, a point reinforced by Kline (2016). Furthermore, this study demonstrates the effectiveness of Rasch analysis as a tool in test development and validation. Linacre (2002) notes that Rasch's analysis offers a powerful method for examining test dimensionality and identifying potential issues with test items. The current findings support the application of Rasch analysis in the test development and validation process, highlighting its importance as emphasized by Wright and Masters (2018). These results collectively underscore the value of rigorous psychometric analysis in educational assessment. They demonstrate how such analyses can enhance the quality and interpretability of test scores, ultimately leading to more accurate and meaningful assessments of student knowledge and skills. The study's findings thus contribute to the broader field of educational measurement by reinforcing best practices in test

development and validation. Overall, the findings of this study provide strong evidence for the one- dimension nature of the GSP122 ICT test, supporting its use as a measure of ICT Proficiency.

Model-Data fit statistics

The analysis of the GSP122 ICT test's model data fit statistics has yielded positive results, demonstrating that the test meets the expectations of the Rasch model assumptions. Both the infit and outfit MNSQ values for all test items fall within the acceptable range of 0.6 to 1.4, as established by Bond and Fox (2015). Additionally, the point measure correlation values for all items were found to be positive and non-zero. These findings provide strong evidence for the one-dimension nature and construct validity of the GSP122 test. This alignment with Rasch model expectations is significant, as Baghaei (2008) asserts that such conformity indicates the scale's relevant validity. Furthermore, Bond and Fox (2001) propose that when items fit Rasch's expectations, it confirms both the one- dimension nature and construct validity of the scale. This suggests that each item in the GSP122 ICT test contributes to measuring a single, intended construct.

In essence, these results support the conclusion that the GSP122 ICT test is a psychometrically sound instrument, measuring a unified construct as intended and demonstrating strong construct validity. This provides confidence in the test's ability to accurately assess the knowledge or skills it was designed to measure among the student population at Jigawa State University. Moreover, Linacre (2010) showed that items contributing meaningfully to the measurement of a single construct are those whose point measure correlation value is positive and not equal to zero. Thus, the GSP122 test measures one construct namely, ICT Proficiency.

Furthermore, since the GSP122 test items showed acceptable fit statistics, meaning they fulfilled the requirements of the Rasch model (Wu & Chen, 2020), the outfit and infit mean square values ranged from 1.37 to .96, indicating that the items were assessing the intended construct, that is, ICT knowledge, and that the items are also functioning as intended (Khalifa & Liao, 2021). This result is in line with other research that examined test item performance using the Rasch

measurement model. For instance, Zhang and Wang's (2020) study discovered that the Rasch measurement model could detect items that were not operating as intended and that eliminating such items enhanced the model's overall fit. Likewise, a study by Alsubaie and Alquraini (2022) found that the Rasch measurement model was able to detect items that were too easy or too difficult, and that the adjustment of these items improved the overall reliability of the test. The acceptable Model data fit statistics in this study indicate that the GSP122 ICT test items are functioning as intended and measuring the intended construct of ICT Proficiency. This provides evidence for the validity and reliability of the test and supports the use of the Rasch model as a tool for test development and evaluation (Lin et al., 2022). Consequently, the application of Baghaei's framework and the use of Rasch analysis have provided robust evidence for the construct validity of the ICT proficiency test. While the majority of items demonstrate good fit, the identification of a few potentially problematic items offers opportunities for further refinement. This ongoing process of evaluation and improvement is essential in maintaining the validity and relevance of educational assessments, particularly in rapidly evolving fields like information and communication technology.

Conclusion

This study set out to evaluate the dimensionality and model-data fit statistics of an Information and Communication Technology (ICT) test administered at a public university in Nigeria using the Rasch model. The findings provide valuable insights into the psychometric properties of this assessment tool, contributing to the broader discourse on ICT proficiency measurement in higher education, particularly in developing countries. The findings also support its use as a valid and reliable tool for assessing students' technological competencies across various academic disciplines. As Nigeria and other developing nations strive to harness the power of ICT for national development, the importance of accurate and fair assessment of digital competencies cannot be overstated. This research contributes to the growing body of literature on ICT education in developing countries and sets a precedent for rigorous evaluation of assessment tools in this context. By ensuring the quality of ICT proficiency measures, we can better support students in

developing the technological skills needed for success in the 21st-century global economy. Moving forward, continued research and refinement of ICT assessment tools, coupled with responsive educational policies and practices, will be crucial in preparing a digitally literate workforce capable of driving innovation and economic growth in Nigeria and beyond. As we navigate the challenges and opportunities of the digital age, high-quality assessments of ICT proficiency will remain an essential component of effective educational strategies, guiding our efforts to bridge the digital divide and empower the next generation of technological leaders.

Recommendations

Based on the findings of this study, it is recommended that future revisions of the ICT test incorporate additional items to enhance content coverage and ensure a more comprehensive assessment of students' digital competencies. Regular psychometric evaluations using the Rasch model should also be conducted to maintain the test's validity, reliability, and fairness across diverse student populations. Furthermore, integrating ICT proficiency assessments into higher education curricula as standardized measures will help track students' progress and identify areas requiring intervention. The adoption of technology-enhanced assessments, such as computer-based and adaptive testing, can further improve the precision of ICT proficiency measurement. Implementing these recommendations will enhance ICT proficiency assessment and contribute to the development of a digitally skilled workforce capable of driving innovation and economic growth in Nigeria and other developing nations.

Implication of the study for Theory and Practice

The findings of this study, which examined the dimensionality and model-data fit statistics of an ICT proficiency test using the Rasch model, have significant implications for both theoretical understanding and practical applications in the field of educational assessment, particularly in the context of ICT education in developing countries.

From a theoretical perspective, the strong evidence for one- dimension nature in the ICT test contributes substantially to the discourse on the nature of ICT proficiency. It suggests that

despite the diverse range of skills and knowledge areas encompassed by ICT literacy, these competencies may be underpinned by a single, cohesive construct. This finding aligns with and extends previous theoretical frameworks that conceptualize ICT literacy as a unified skill set (e.g., Covello, 2010). Furthermore, the good fit of most items to the Rasch model provides empirical support for the construct validity of ICT proficiency as measured by this test. This strengthens the theoretical foundation for assessing ICT skills and supports the idea that these skills can be reliably measured across different levels of ability.

From a practical standpoint, the validated one- dimension nature of the test supports the use of total scores as meaningful measures of overall ICT proficiency. Practitioners can confidently use these scores for various educational decisions, knowing they represent a coherent construct. This has immediate implications for assessment practices in Nigerian universities and could lead to more standardized approaches to evaluating students' digital competencies. The validated psychometric properties of the test provide a benchmark for quality assurance in ICT assessment. Other institutions in Nigeria and beyond can use this study as a model for evaluating their own assessment tools, potentially leading to more rigorous and standardized approaches to measuring digital literacy across the higher education sector.

From a policy perspective, these findings can inform national strategies for digital literacy education. The validated assessment tool provides a means to measure progress towards national ICT proficiency goals, aligning with Nigeria's broader objectives for technological advancement and economic development. Policymakers can use these insights to develop evidence-based strategies for enhancing digital skills across the education system. The comprehensive coverage of ICT skills in the test can help ensure that higher education institutions are preparing students with the digital competencies required by employers, potentially enhancing graduate employability. This alignment between academic assessment and industry needs is crucial for Nigeria's economic growth and competitiveness in the global digital economy.

Limitations and Future Directions

While this study provides robust evidence for the psychometric quality of the ICT test, several limitations should be acknowledged. Firstly, the cross-sectional nature of the research design offers a snapshot of the test's performance at a single point in time. Longitudinal studies could provide valuable insights into the stability of the test's psychometric properties over time and its sensitivity to changes in students' ICT proficiency levels. Secondly, while the sample size was adequate for Rasch analysis, larger-scale studies involving multiple institutions could further validate these findings and explore potential variations across different higher education contexts in Nigeria.

Thirdly, this study focused primarily on the internal structure and functioning of the test items. Future research could complement these findings by examining the test's external validity, perhaps by correlating test scores with other measures of ICT proficiency or real-world performance in technology-related tasks.

Several avenues for future research emerge from this study. One promising direction would be to investigate the relationship between students' performance on this ICT test and their academic achievement across different disciplines. Such research could provide insights into the role of ICT proficiency in supporting learning outcomes in various fields of study. Another important area for future inquiry is the exploration of potential gender differences in test performance and item functioning. While this study focused on disciplinary differences, understanding any gender-related disparities in ICT proficiency could inform efforts to promote gender equity in technological education.

Acknowledgment

academic environments fostered by both institutions, which greatly facilitated our research endeavors, and their support is deeply appreciated.

Disclosure Statement

References

Adebisi, T. A., Oyeleke, O., & Adeleke, A. G. (2018). Digital literacy skills and MOOC participation among lecturers in a Nigerian private university. *Library Philosophy and Practice*, 1-18. https://core.ac.uk/download/pdf/188131266.pdf

Alsubaie, M. A., & Alquraini, A. A. (2022). Using the Rasch model to evaluate the quality of test items in a computer programming course. *Journal of Educational Computing Research, 61*(4), 419–433.

Adedokun-Shittu, N. A., & Shittu, A. J. K. (2015). Assessing the impacts of ICT deployment in teaching and learning in higher education: Using ICT impact assessment model. *Journal of Applied Research in Higher Education, 7*(2), 180-193. https://doi.org/10.1108/JARHE-02-2013-0012

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). Routledge.

Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch analysis in the human sciences*. Springer.

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

Etikan, I., & Bala, K. (2017). Sampling and sampling methods. *Biometrics & Biostatistics International Journal, 5*(6), 00149. https://doi.org/10.15406/bbij.2017.05.00149.

Eakman, A. M. (2012). The measurement characteristics of the engagement in meaningful activities surveys in an age-diverse samples. *The American Journal of Occupational Therapy, 66*(2), 20–29. https://doi.org/10.5014/ajot.2012.001867.

Fisher, W.P. (2007). Rating scale instrument quality criteria. *Rasch measurement transactions, 21*(1), 1095. http://www.rasch.org/rmt/rmt211a.htm

Hamad, A. A. (2021). The Psychometric Properties of Measurement of the Mathematics Teachers' Professional Identity in Saudi Arabia (Unpublished Doctoral dissertation, USM). http://eprints.usm.my/52981/1/ALGHAMDI%20HAMAD%20ABDULLAH%20J%20-%20TESIS.pdf%20cut.pdf

Khalifa, B. S., & Liao, S. Y. (2021). Applying the Rasch model to evaluate the performance of test

items in a mathematics assessment. *International Journal of Assessment and Evaluation, 28*(1), 1–15. DOI: 10.7763/IJSSH.2012.V2.104

Kline, P. (2016). *A handbook of test development: Guide for professionals in education, psychology, and business*. Sage Publications.

Lin, H. C., Wu, C. H., & Chen, S. Y. (2022). Using the Rasch model to develop and evaluate a reading comprehension test. *Reading and Writing, 35*(3), 537–555.

Levin, K. A. (2006). Study design III: Cross-sectional studies. *Evidence-Based Dentistry*, 7(1), 24-25. https://doi.org/10.1038/sj.ebd.6400375

Linacre, J. M. (1994). Sample size and item calibration stability. *Rasch Measurement Transactions*, 7(4), 328.

Linacre, J. M. (2021). *Winsteps® Rasch measurement computer program user's guide*. Winsteps.com.

Linacre, J. M. (2006). *A user's guide to WINSTEPS/ MINISTEPS*: A Rasch model computer programs. Chicago, USA: Winsteps.com.

Linacre, John M. (2007). A user's guide to WINSTEPS: *Ministep Rasch Model Computer Programs*. MESA Press.

Linacre, G. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement, 3*(1), 85-106. https://pubmed.ncbi.nlm.nih.gov/11997586/

Linacre, G. M. (2010). *A User's Guide to winsteps*. Program Manual 3.71.0.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, *50*(9), 741-749. https://doi.org/10.1037/0003-066X.50.9.741.

Olaniyi, O. A, & Ismaila, K. O. (2016). Information and communication technologies (ICTs) usage and household food security status of maize crop farmers in Kwara State, Nigeria. *Library Philosophy and Practice*, 1-20.

Olelewe, C. J., & Agomuo, E. E. (2016). Effects of B-learning and F2F learning environments on students' achievement in QBASIC programming. *Computers & Education*, 103, 76-86.

https://doi.org/10.1016/j.compedu.2016.09.012.

Paek, I., & Wilson, M. (2018). A Rasch analysis of the Test of English as a Foreign Language (TOEFL). *Language Testing, 35*(2), 147-164.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Nielsen & Lydiche.

Reckase, D. (1979). Unifactor latent traits models applied to multifactor tests: results and implications. *Journal of Educational Statistics, 4*(3), 207–230. https://doi.org/10.3102/10769986004003207.

Smith, P. (2017). Validity theory and the Rasch model. Measurement: *Interdisciplinary Research and Perspectives, 15*(1), 34-46. https://doi.org/10.1080/15366367.2017.1285137

Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. *Rasch Measurement Transactions*, *8*(3), 370.

Wright, B. D., & Masters, G. N. (2018). *Rating scale analysis: Rasch measurement*. SAGE Publications.

Wu, Y. W., & Chen, S. Y. (2020). Using the Rasch model to evaluate the performance of test items in a science assessment. *Journal of Science, Education, and Technology, 29*(2), 249–262.

Yang, P., et al. (2020). A Rasch analysis of the Chinese version of the Beck Depression Inventory-II. *Journal of Clinical Psychology, 76*(1), 35-47.

Zwick, R., Thayer, D. T., & Lewis, C. (1999). An empirical Bayes approach to Mantel-Haenszel DIF analysis. *Journal of Educational Measurement*, *36*(1), 1-28. https://doi.org/10.1111/j.1745-3984.1999.tb00543.x

Zhang, Y., & Wang, W. C. (2020). Applying the Rasch model to evaluate the quality of test items in a language proficiency test. *Language Testing, 37*(3), 349–366. DOI: 10.21831/cp.v38i1.22750