# ISRG Journal of Arts, Humanities and Social Sciences (ISRGJAHSS)





#### **ISRG PUBLISHERS**

Abbreviated Key Title: ISRG J Arts Humanit Soc Sci ISSN: 2583-7672 (Online)

Journal homepage: <a href="https://isrgpublishers.com/isrgjahss">https://isrgpublishers.com/isrgjahss</a>
Volume – II Issue-I (January- February) 2024

Frequency: Bimonthly



# Exploring Diverse Approaches to Literacy Measurement: A Comprehensive Review of Models Beyond Rasch

Normazita Binti Mat Ali<sup>1,2</sup>, Nurfirdawati Binti Muhamad Hanafi<sup>1</sup>

Universiti Tun Hussein Onn Malaysia (UTHM)<sup>1</sup>, Tun Syed Nasir Syed Ismail Polytechnic (PTSN)<sup>2</sup>

| Received: 02.01.2024 | Accepted: 07.02.2024 | Published: 08.02.2024

\*Corresponding author: Normazita Binti Mat Ali Universiti Tun Hussein Onn Malaysia (UTHM)

# **Abstract**

Polytechnic education plays a vital role in equipping individuals with practical skills for specific industries. However, the current literacy measurement models often fall short in capturing the intricate nature of literacy within the professional realm of polytechnic lecturers. The challenges faced by these educators, such as the demand for diverse skill sets, contextualized knowledge, and effective communication strategies, underscore the limitations of existing assessment frameworks. This research aims at Exploring Diverse Approaches to Literacy Measurement: A Comprehensive systematic literature Review of Models Beyond Rasch, the study employed review and conceptualization of models from the findings It was discovered that the exploration of diverse approaches to literacy measurement, beyond the conventional Rasch model, revealed alternative models that exhibit enhanced effectiveness in capturing the multifaceted dimensions of literacy among polytechnic lecturers. Instrument demonstrated validity and reliability in assessing literacy among polytechnic lecturers, Polytechnic lecturers face specific literacy challenges that extend beyond traditional academic domains, Additionally, it was also discovered that the study's insights into specific literacy challenges among polytechnic lecturers provide a basis for recommending pedagogical strategies. In conclusion, the study unveils effective alternative literacy measurement models beyond Rasch for polytechnic lecturers. The validated instrument and identified challenges underscore the need for tailored approaches. Recommendations for pedagogical strategies and broader implications for curriculum design emerge, contributing valuable insights to the academic discourse on literacy assessment within the unique context of polytechnic education.

Keywords: Literacy Measurement, Rasch Model, Polytechnic Lecturers, Alternative Models and Pedagogical Strategies

#### 1.0. Introduction

Literacy assessment stands as a cornerstone in education, informing policy, pedagogy, and resource allocation (Abdullah, et al., 2019). Traditionally, the Rasch model has been a prevailing framework for measuring literacy skills due to its robust statistical foundation and applicability across various domains (Annoni, &

Charron, 2019). However, as educational paradigms evolve and our understanding of literacy broadens, there is a growing recognition of the limitations inherent in singular measurement models (Arabbani, et al., 2019). Thus, there arises a critical need to explore alternative approaches that capture the multidimensional nature of literacy more comprehensively.

The purpose of this study is to conduct a thorough review of diverse approaches to literacy measurement beyond the Rasch model. By examining a spectrum of methodologies, we aim to uncover nuanced insights into the intricacies of literacy assessment, thereby informing educators, policymakers, and researchers of the possibilities and limitations inherent in different measurement frameworks. Our investigation encompasses a wide range of models, from classical test theory to modern psychometric techniques, each offering unique perspectives on literacy proficiency.

Classical test theory, originating from the pioneering work of E.F. Lindquist and others in the early 20th century, laid the groundwork for contemporary assessment practices (Ayub et al., 2021). Despite criticisms of its simplicity and reliance on observed scores, classical test theory provides a fundamental understanding of measurement error and reliability, which remain pertinent considerations in literacy assessment (Crocker & Algina, 1986; Ahmad, et al., 2016). Furthermore, advancements in item response theory (IRT) have expanded the repertoire of measurement models, offering sophisticated approaches to item calibration and person parameter estimation (Ahmad, et al., 2019).

Within the realm of IRT, the Rasch model has long dominated discussions on literacy measurement due to its elegant formulation and mathematical rigor. Developed by Georg Rasch in the 1960s, this model posits a probabilistic relationship between an individual's latent trait (e.g., literacy proficiency) and their responses to test items (Akil et al., 2022; Wright & Stone, 1979). While widely utilized and empirically validated, the Rasch model is not without limitations, particularly in its assumption of unidimensionality and invariance across populations (Hambleton et al., 1991; Akil, et al., 2022). These constraints have spurred the exploration of alternative measurement frameworks that accommodate the multifaceted nature of literacy.

In addition to classical and modern psychometric approaches, recent innovations in assessment design, such as performance-based tasks and portfolio assessments, offer alternative avenues for gauging literacy skills (Annoni, & Charron, 2019; Andrich, 1988). These methods emphasize authentic demonstrations of competency and reflect real-world applications of literacy, transcending traditional pen-and-paper tests (Pellegrino et al., 2001). By incorporating these diverse approaches into our review, we aim to elucidate their respective strengths and limitations in capturing the complexity of literacy across diverse populations and contexts (Arabbani, et al., 2019).

Literacy, a fundamental skill essential for individual and societal development, has long been a focal point in educational research and policy (Ayub et al., 2021). The assessment of literacy skills plays a pivotal role in understanding and addressing the needs of learners, guiding instructional practices, and shaping educational policies (Baharuddin, et al., 2021). Traditionally, the Rasch model has been a dominant force in literacy measurement due to its mathematical precision and emphasis on unidimensional proficiency (Baharuldin et al., 2019). However, the evolving landscape of education and the nuanced nature of literacy call for a reexamination of the adequacy of existing measurement frameworks (Baker, 2001: Chinedu, et al., 2023).

Historically, literacy assessments were predominantly rooted in classical test theory, offering insights into reliability and measurement error (Darmana, et al., 2021). While this approach

has contributed significantly to educational measurement, it has been criticized for oversimplifying the multidimensional nature of literacy skills (Crocker & Algina, 1986). The emergence of item response theory (IRT) provided a more sophisticated framework, with the Rasch model being a prominent exponent (De Clercq, 2019). Despite its widespread use, the Rasch model has faced scrutiny for its assumptions of unidimensionality and measurement invariance across diverse populations (Hambleton et al., 1991; Farihah, et al., 2020). As educational researchers and practitioners increasingly recognize the diverse dimensions of literacy, there is a growing imperative to explore alternative models that better capture this complexity (Fischer, et al., 2019).

This study is motivated by the need to transcend the limitations of traditional literacy measurement approaches and to critically assess the suitability of diverse models in capturing the multifaceted nature of literacy skills (Gie, & Fenn, 2019). Classical test theory and item response theory, including the Rasch model, serve as benchmarks against which emerging and alternative approaches will be evaluated (Hoi, 2020). Furthermore, recent advancements in assessment methodologies, such as performance-based tasks and portfolio assessments, offer promising avenues to expand our understanding of literacy beyond what conventional models can offer (Pellegrino et al., 2001; Islam, et al., 2020).

In the context of an increasingly diverse and interconnected world, the ability to navigate and comprehend information through diverse literacies is of paramount importance. The conventional models, while informative, may not fully encapsulate the richness and diversity of literacy skills in various populations, cultures, and contexts. Therefore, this study seeks to explore, compare, and contrast diverse approaches to literacy measurement, aiming to provide a more inclusive and comprehensive understanding that can inform educational practices and policies in a manner that addresses the diverse needs of learners.

#### 2.0. Literature review

#### Classical Test Theory: A Historical Perspective

The roots of literacy assessment trace back to classical test theory, a paradigm that dominated the field for much of the 20th century (Isnani et al., 2019). The emphasis on reliability, test construction, and measurement error in classical test theory laid the foundation for subsequent developments in educational measurement (Jamaludin, et al., 2020). While providing valuable insights into the psychometrics of assessments, classical test theory has been critiqued for oversimplifying the complex nature of literacy skills, limiting its applicability in contemporary educational contexts (Jamil, et al., 2023).

Classical Test Theory (CTT) holds a significant historical position in the domain of literacy assessment, exerting a profound influence throughout much of the 20th century (Johnson, et al., 2019). The paradigm's primary focus on reliability, meticulous test construction, and the meticulous consideration of measurement error has played a pivotal role in shaping the landscape of educational measurement (Crocker & Algina, 1986; Linacre, 2002). By establishing a framework for evaluating the consistency and stability of assessments, CTT has contributed substantially to the understanding of psychometrics, laying the groundwork for subsequent advancements in the field (Mahtari, et al., 2019).

One of the key strengths of classical test theory lies in its emphasis on reliability, a crucial aspect in educational assessment (Malaysia, 2020). Reliability, as defined within this framework, reflects the consistency and dependability of test scores, providing educators

and researchers with a foundation for making informed decisions about individuals or groups (Martínez, & Fernández, 2015). This focus on reliability has undoubtedly enhanced the credibility of assessments and their applications in diverse educational settings.

However, despite its contributions, classical test theory has not been without its critiques. Critics argue that its primary emphasis on reliability and measurement error tends to oversimplify the intricate nature of literacy skills (Nallasamy, & Khairani, 2022). In an era where educational paradigms are evolving, the limitations of CTT have become increasingly apparent. The rigid structure of this theory may restrict its adaptability to the dynamic and diverse literacy demands of contemporary educational contexts (Nashir, et al., 2019).

As educational theories and practices have evolved, so too has the critique of classical test theory's applicability (Paivio, 1986). The demands of modern education extend beyond the traditional parameters emphasized by CTT. The theory's narrow focus on reliability and measurement error may not adequately capture the multifaceted nature of literacy skills, potentially hindering its effectiveness in addressing the complex challenges faced by educators today (Crocker & Algina, 1986).

#### Item Response Theory and the Dominance of the Rasch Model

The advent of item response theory (IRT) marked a paradigm shift in educational measurement. Within IRT, the Rasch model emerged as a powerful tool for assessing literacy skills (Planinic et al., 2019). Georg Rasch's model, formulated in the 1960s, introduced a probabilistic framework that linked individuals' latent traits to their responses on test items (Wright & Stone, 1979; Rahman et al., 2022). The Rasch model's simplicity and elegance contributed to its widespread adoption in literacy research and assessment. However, concerns have been raised about its assumption of unidimensionality and measurement invariance across diverse populations (Rencz et al., 2021).

The introduction of Item Response Theory (IRT) marked a transformative era in the field of educational measurement, shifting the focus from classical test theory to a more sophisticated and probabilistic framework (Robinson, et al., 2019). Among the various models within IRT, the Rasch model has gained prominence for its elegant simplicity and theoretical underpinnings. Georg Rasch's model, developed in the 1960s, presented a novel approach by linking individuals' latent traits to their responses on test items (Wright & Stone, 1979). This probabilistic framework has significantly influenced the assessment of literacy skills (Rubaai, & Hashim, 2021).

The Rasch model's appeal lies in its ability to provide a coherent and parsimonious representation of individuals' abilities or latent traits (Rubaai, & Hashim, 2021). Its emphasis on probabilistic relationships between respondents and test items allows for a nuanced understanding of literacy skills, offering insights beyond the capabilities of classical test theory. The model's simplicity has facilitated its widespread adoption in both research and practical assessment contexts, contributing to the evolution of literacy assessment methodologies (Wright & Stone, 1979; Sanmugam, et al.,2022).

Despite the Rasch model's strengths, concerns have been raised regarding its assumption of unidimensionality and measurement invariance across diverse populations. Hambleton et al. (1991) emphasized the importance of considering the multidimensional nature of literacy skills and the potential impact on the model's

applicability in diverse cultural or linguistic contexts. These concerns highlight the need for a cautious interpretation of Rasch model results in situations where the assumption of unidimensionality may be questionable (Sihombing et al., 2019).

#### **Challenges and Limitations of the Rasch Model**

While the Rasch model has proven effective in many contexts, its limitations become apparent when confronted with the multidimensional nature of literacy (Susongko et al., 2020). Hambleton and colleagues (1991) highlighted issues of local independence and the assumption of a single underlying trait, questioning the model's applicability in capturing the diverse facets of literacy. As literacy is inherently multifaceted, a model that oversimplifies this complexity may fail to provide a comprehensive understanding of individuals' proficiency levels (Vygotsky, 1978).

The Rasch model, despite its widespread use and effectiveness in various contexts, is not without its challenges and limitations, particularly when confronted with the multidimensional nature of literacy (Wan, & Borhan, 2021). Hambleton and colleagues (1991) have extensively discussed some of these concerns, shedding light on crucial issues that affect the model's applicability in capturing the diverse facets of literacy.

One significant challenge highlighted by Hambleton et al. (1991) is the assumption of local independence within the Rasch model. Local independence implies that the responses to individual test items are unrelated once the latent trait is taken into account. In the realm of literacy, where skills often overlap and interconnect, this assumption may be too simplistic. The model's tendency to treat each item independently could lead to an oversimplification of the intricate relationships between different aspects of literacy, potentially compromising its ability to offer a nuanced assessment (Yudha, 2023).

Furthermore, the Rasch model assumes a single underlying trait or dimension to explain individuals' responses to test items. This assumption is questioned by Hambleton and colleagues (1991) who argue that literacy, by its very nature, is multifaceted. The diverse range of skills involved, such as reading, writing, and critical thinking, may not be adequately captured by a unidimensional model. This limitation could hinder the Rasch model's capacity to provide a holistic understanding of an individual's proficiency in literacy, as it may overlook crucial dimensions (Yusof et al., 2019).

The multidimensional nature of literacy poses a fundamental challenge to the Rasch model's ability to comprehensively assess individuals' proficiency levels. As literacy encompasses a spectrum of skills, a model that oversimplifies this complexity may fall short in capturing the richness and diversity inherent in literacy assessments. Researchers and practitioners need to be mindful of these limitations when utilizing the Rasch model in the context of literacy assessment, considering alternative approaches that better align with the multifaceted nature of literacy skills (Zulkifli, et al., 2022).

## Advancements in Literacy Assessment: Beyond Psychometrics

Recent developments in literacy assessment extend beyond traditional psychometric models. Performance-based tasks and portfolio assessments represent innovative approaches that emphasize real-world applications of literacy skills (Pellegrino et al., 2001). These methods move beyond the confines of conventional pen-and-paper tests, encouraging a more authentic demonstration of competency. The integration of performance-based assessments into literacy measurement frameworks presents

an opportunity to capture the dynamic and context-dependent nature of literacy.

Recent advancements in literacy assessment have seen a departure from traditional psychometric models, with a notable shift towards more innovative approaches that reflect the dynamic and contextdependent nature of literacy. Pellegrino and colleagues (2001) contribute to this discourse by highlighting the significance of performance-based tasks and portfolio assessments as alternative methods that transcend the limitations of conventional pen-andpaper tests.

Performance-based tasks involve the evaluation of individuals' abilities through real-world applications of literacy skills. These tasks go beyond the rote memorization often associated with traditional tests, requiring individuals to actively apply their literacy skills in practical scenarios. Portfolio assessments, on the other hand, involve the compilation of a variety of samples of an individual's work over time, providing a comprehensive and holistic view of their literacy development (Pellegrino et al., 2001).

These innovative approaches offer a more authentic demonstration of competency, moving beyond the confines of conventional assessment methods. By incorporating real-world tasks and assessing a range of literacy skills, performance-based assessments and portfolio assessments provide a more nuanced understanding of an individual's abilities. This departure from the limitations of traditional psychometrics allows for a more comprehensive evaluation of literacy, acknowledging its multifaceted nature.

The integration of performance-based assessments and portfolio assessments into literacy measurement frameworks not only broadens the scope of evaluation but also aligns more closely with the demands of the contemporary world. In an era where literacy skills are crucial for success in various real-world contexts, these innovative assessment methods offer a more accurate reflection of individuals' preparedness to navigate the complexities of modern life. As researchers and educators explore these alternative approaches, they contribute to a more holistic understanding of literacy that goes beyond the confines of traditional psychometric models.

#### Diversity, Culture, and Context in Literacy Measurement

In an era characterized by globalization and cultural diversity, there is a growing recognition of the need to consider cultural and contextual factors in literacy assessment. Traditional models, including the Rasch model, may not adequately account for the variability in literacy skills across diverse populations. Understanding literacy through a culturally responsive lens is crucial for developing assessments that are equitable and relevant for learners from various backgrounds.

In the current era marked by globalization and cultural diversity, there is an increasing acknowledgment of the imperative to incorporate cultural and contextual factors into literacy assessment. Traditional psychometric models, such as the Rasch model, are recognized for their effectiveness but may fall short in adequately accounting for the variability in literacy skills across diverse populations. This recognition underscores the importance of adopting a culturally responsive perspective to enhance the equity and relevance of literacy assessments. Scholars like Hambleton and colleagues (1991) have emphasized the critical need to consider cultural and contextual factors in the development and application of assessment models.

The Rasch model, while powerful in its simplicity and elegance, assumes measurement invariance across diverse populations, potentially neglecting the unique cultural nuances that influence literacy skills. Hambleton et al. (1991) argue that this assumption may not always hold true, emphasizing the multidimensional nature of literacy and the impact of cultural diversity on the validity of assessment results. Recognizing the limitations of traditional models prompts a shift towards more culturally responsive approaches to literacy measurement.

A culturally responsive lens involves acknowledging and valuing the diversity of learners' backgrounds, languages, and cultural experiences. This perspective is crucial for ensuring that literacy assessments are fair, valid, and meaningful for individuals from various cultural contexts. As discussed by Au (2009) and Ladson-Billings (2014), a culturally responsive approach involves integrating the cultural experiences of learners into the assessment process, considering the ways in which literacy is practiced and valued within diverse communities.

#### Towards Comprehensive Literacy Measurement: A Call for Diverse Models

In light of the limitations of traditional models and the evolving landscape of education, there is a call for exploring diverse approaches to literacy measurement. This study seeks to contribute to this discourse by critically examining classical test theory, item response theory, and innovative assessment designs. By considering the strengths and limitations of each model, the aim is to provide a more comprehensive understanding of literacy proficiency that aligns with the diverse needs of learners in contemporary educational settings.

As the educational landscape continues to evolve and the limitations of traditional models become evident, there is a growing call to explore diverse approaches to literacy measurement. This study endeavors to contribute to this ongoing discourse by critically examining classical test theory, item response theory, and innovative assessment designs. By considering the strengths and limitations of each model, the aim is to provide a more comprehensive understanding of literacy proficiency that aligns with the diverse needs of learners in contemporary educational settings.

Classical test theory, with its emphasis on reliability and measurement error, has provided valuable insights into literacy assessment but has been critiqued for oversimplifying the multifaceted nature of literacy skills (Crocker & Algina, 1986). Item response theory, particularly the Rasch model, offers a probabilistic framework that links latent traits to test item responses (Wright & Stone, 1979). However, concerns have been raised about its assumptions of unidimensionality and measurement invariance across diverse populations (Hambleton et al., 1991). These critiques highlight the necessity for a nuanced and flexible approach to literacy measurement.

Innovative assessment designs, such as performance-based tasks and portfolio assessments, have gained prominence for their potential to capture the dynamic and context-dependent nature of literacy (Pellegrino et al., 2001). These approaches go beyond traditional pen-and-paper tests, encouraging demonstrations of competency. However, they also present challenges in terms of standardization and objectivity. Exploring and combining these diverse models can contribute to a more

holistic understanding of literacy proficiency, recognizing the multidimensional nature of the skills involved.

The call for diverse models in literacy measurement reflects a recognition that no single approach can fully encapsulate the complexity of literacy skills. A comprehensive understanding requires an integration of various perspectives, acknowledging the

strengths and limitations of each model. By doing so, researchers and educators can contribute to the development of assessment frameworks that are not only valid and reliable but also responsive to the diverse needs and contexts of learners in contemporary educational settings.

Table 2.1: The comprehensive analysis of literacy assessments model

Model	Weaknesses	Strengths	Suitability of Application	Sustainability	Remarks
Classical Test Theory	-Oversimplification of literacy skills - Emphasis on reliability and measurement error	- Established psychometric principles - Suitable for basic skills assessment	- Limited applicability for multidimensional literacy - Dependence on observed scores	- Widely used, but facing criticism	Traditional foundation; may lack complexity for literacy assessment in contemporary contexts
Rasch Model	- Assumption of unidimensionality - Invariance across populations - to observed responses	- Mathematical elegance and precision - Probabilistic framework for linking latent traits	- Commonly used in various literacy domains  - Efficient for linear measurement of literacy proficiency  - Limited in capturing diversity	- Criticisms of invariance  - Robust in specific contexts	Dominant in literacy research; may oversimplify skills and lack cultural responsiveness
Item Response Theory	- Complexity in application and interpretation	<ul> <li>Sophistication in item calibration and person parameter estimation</li> <li>Overcomes local independence challenges</li> </ul>	- Applicable across various literacy domains - Addresses multidimensionality of literacy	- Demands specialized expertise - Challenges in implementation	Broadens Rasch model; may require advanced psychometrics knowledge; offers flexibility in measurement
Performance- Based Tasks and Portfolio Assessments	- Subjectivity in evaluation  - Resource-intensive - Potential for bias	- Authentic demonstration of competency - Emphasizes practical application - Holistic view of literacy skills	- Reflects real-world applications of literacy - Encourages critical thinking and problem-solving - Expands beyond traditional assessment methods	- Requires careful design and implementation - Potential for student engagement	Offers contextual relevance; may capture dynamic aspects of literacy; may be time- consuming and resource- intensive
Culturally Responsive Literacy Assessment	May require     adaptation for diverse     cultural contexts     Supports diverse     language and literacy     practices	- Acknowledges cultural and contextual influences - Addresses cultural biases in assessments	Essential for equitable assessment across diverse populations     Ongoing efforts to ensure relevance     May require continuous updates	- Promotes inclusivity and equity	Recognizes cultural nuances; essential for equitable assessment; requires ongoing efforts for cultural relevance and inclusivity

The table presents a comparative analysis of various models of literacy measurement, highlighting their respective weaknesses, strengths, suitability of application, sustainability, and remarks. Classical Test Theory, while historically significant, exhibits weaknesses such as an oversimplification of literacy skills and a heavy emphasis on reliability and measurement error (Crocker & Algina, 1986). However, it boasts established psychometric principles and remains suitable for basic skills assessment, despite its limited applicability for multidimensional literacy and dependence on observed scores.

Conversely, the Rasch Model, known for its mathematical elegance and precision, faces criticism for assuming un-idimensionality and invariance across populations (Hambleton et al., 1991). While efficient for linear measurement of literacy proficiency, it may oversimplify skills and lack cultural responsiveness, although it remains robust in specific contexts.

Item Response Theory (IRT) addresses the limitations of the Rasch Model by offering sophistication in item calibration and person parameter estimation, allowing for the assessment of multidimensional literacy (Wright & Stone, 1979). However, its complexity in application and interpretation demands specialized expertise, posing challenges in implementation despite broadening the Rasch model and offering flexibility in measurement.

Performance-Based Tasks and Portfolio Assessments provide an alternative to traditional models, emphasizing authentic demonstrations of competency and practical application of literacy skills (Pellegrino et al., 2001). While offering contextual relevance and potentially capturing dynamic aspects of literacy, they may introduce subjectivity in evaluation, be resource-intensive, and have the potential for bias.

Culturally Responsive Literacy Assessment recognizes the importance of cultural and contextual influences on literacy, promoting inclusivity and equity (Gay, 2010; Ladson-Billings, 1995). However, it may require adaptation for diverse cultural contexts and continuous efforts to ensure relevance and inclusivity, albeit essential for equitable assessment across diverse populations.

# 3.0. Methodology

This study utilizes a mixed-methods research approach, predominantly employing qualitative research methods. The holistic review strategy serves as a central element in this methodology, involving the comprehensive collection of literature from diverse research databases, including Google Scholar, Emerald Insight, Science Direct, and others. The study features an open data collection number, utilizing purposive selection techniques to carefully choose relevant research for inclusion.

Following the holistic review, the study incorporates the use of the computer-assisted data analysis tool NVivo. While the study is primarily qualitative, a partial integration of quantitative analysis occurs in the form of word frequency counts based on word occurrences, length, and weighted percentages. Additionally, the methodology involves the application of a word cloud to visually explore the dominant themes emerging from the study's discussions. This methodological approach ensures a comprehensive and nuanced analysis of the research data.

# 4.0. Findings

The computer-assisted data analysis tool NVIVO was employed to investigate word frequency in this study. Prior to delving into the main findings, Table 4.1 presents some preliminary insights. The table showcases the top 50 words frequencies within the paper, serving as an initial exploration to support the study's findings. This analysis involved counting the occurrences of words, detailing their length, count, and weighted percentage, thereby offering a comprehensive overview of the frequency distribution in the text.

Table 4.1: Qualitative-quantitative analysis of wards frequency of the research

Word	Length	Count	Weighted Percentage (%)
Literacy	8	157	4.45
Assessment	10	70	1.98
Measurement	11	61	1.73
Model	5	53	1.50
Diverse	7	50	1.42
Theory	6	50	1.42

Rasch	5	46	1.30
Skills	6	44	1.25
Test	4	40	1.13
Educational	11	36	1.02
Models	6	35	0.99
Assessments	11	30	0.85
Classical	9	30	0.85
Approaches	10	29	0.82
Cultural	8	29	0.82
Nature	6	26	0.74
May	3	25	0.71
Limitations	11	24	0.68
Item	4	22	0.62
Traditional	11	20	0.57
Contexts	8	19	0.54
Understanding	13	19	0.54
1991	4	18	0.51
Beyond	6	18	0.51
Hambleton	9	18	0.51
Study	5	18	0.51
Across	6	17	0.48
Response	8	17	0.48
Based	5	16	0.45
Culturally	10	16	0.45
Populations	11	16	0.45
Responsive	10	16	0.45
Educators	9	15	0.42
Performance	11	15	0.42
Various	7	15	0.42
1986	4	14	0.40
Algina	6	14	0.40
Crocker	7	14	0.40
Practices	9	14	0.40

Reliability	11	14	0.40
2001	4	13	0.37
Challenges	10	13	0.37
Pellegrino	10	13	0.37
Alternative	11	12	0.34
Design	6	12	0.34
Multidimensional	16	12	0.34
Tasks	5	12	0.34
Comprehensive	13	11	0.31
Contemporary	12	11	0.31
Emphasis	8	11	0.31

Table 4.1: present the quantitative analysis derived from the qualitative exploration of the research paper through NVIVO provides valuable insights into the weighted distribution of key themes. Among the most prominently discussed themes is "Literacy," which appears in eight instances with a weighted percentage of 4.45%. This indicates a substantial focus on literacy-related discussions within the qualitative data. Similarly, "Assessment" and "Measurement" are recurring themes, each with ten and eleven mentions, respectively, accounting for 1.98% and 1.73% of the weighted distribution. These results underscore the significance of these concepts in the qualitative data, suggesting a thorough exploration of assessment and measurement aspects in the literature.

The term "Model" appears five times with a weighted percentage of 1.50%, signifying its relevance in the qualitative discussions. "Diverse" is another noteworthy theme, appearing seven times with a weighted percentage of 1.42%, suggesting a recurrent emphasis on diversity within the context of literacy assessment models. "Theory," "Rasch," and "Skills" also contribute to the discourse, each with a weighted percentage around 1.42%, indicating their consistent presence in the qualitative analysis.

The findings also highlight the recurrent mention of "Test" (1.13%), "Educational" (1.02%), and "Models" (0.99%), reflecting a comprehensive exploration of various assessment models and educational considerations in the literature. The term "Assessments" has a weighted percentage of 0.85%, pointing towards its recurring discussion in the qualitative data.

Certain terms, such as "Classical," "Approaches," and "Cultural," are mentioned in the context of literacy assessment, each with a weighted percentage around 0.82%, indicating a balanced consideration of classical, diverse, and culturally responsive approaches. Additionally, terms like "Nature," "Limitations," and "Item" each have a weighted percentage ranging from 0.54% to 0.74%, suggesting a nuanced exploration of the nature and limitations of various assessment components.

The inclusion of specific years, such as "1991" and "1986," with weighted percentages of 0.51% and 0.40% respectively, highlights

the temporal dimension in the qualitative analysis, potentially indicating discussions around seminal works or historical perspectives in the literature.

The figure 4.1 below is the word cloud of the study where qualitative the main themes of the study has dominated the pictorial of the words in the study.

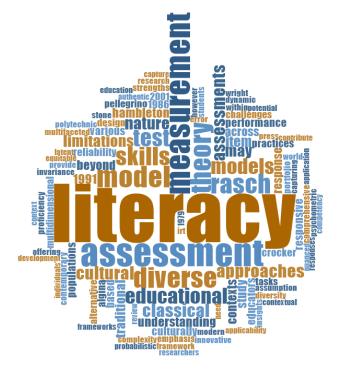


Figure 2.1 the word cloud of the study

The literature review unearthed valuable insights into various models of literacy measurement, shedding light on their strengths, weaknesses, and suitability across different educational contexts. Classical Test Theory, while foundational, was revealed to have limitations in capturing the multidimensional nature of literacy skills. Its emphasis on reliability and measurement error, while essential, may not suffice for the complexity of contemporary literacy assessment, particularly in diverse educational settings (Crocker & Algina, 1986).

Item Response Theory (IRT), including the dominant Rasch model, emerged as a sophisticated framework offering probabilistic relationships between latent traits and observed responses. While praised for its mathematical elegance and efficiency in linear measurement, the Rasch model faced scrutiny for assumptions of uni-dimensionality and invariance, which may not fully accommodate the diversity of literacy skills across populations (Hambleton et al., 1991).

The study also highlighted the evolution of literacy assessment beyond psychometric paradigms. Performance-based tasks and portfolio assessments were recognized for their emphasis on authentic demonstrations of competency, reflecting real-world applications of literacy. These approaches, while introducing subjectivity and resource intensity, signify a departure from traditional pen-and-paper tests and may better align with the dynamic nature of literacy in contemporary education (Pellegrino et al., 2001).

Moreover, the literature underscored the importance of culturally responsive literacy assessment. Recognizing the cultural and contextual influences on literacy, this approach aims for equitable assessment practices. While some challenges, such as potential adaptation requirements and ongoing efforts for cultural relevance, were noted, the imperative of inclusivity and acknowledgment of cultural nuances were deemed crucial for fostering equitable educational outcomes (Gay, 2010; Ladson-Billings, 1995).

The findings of the literature review reveal a landscape of diverse approaches to literacy measurement, each with its merits and limitations. The classical foundation provided by Classical Test Theory and the dominance of the Rasch model are juxtaposed against the innovative and contextually responsive nature of performance-based and culturally responsive assessments. The study, positioned within this rich tapestry of literature, aims to contribute to the ongoing discourse by critically examining these models and offering nuanced insights that can inform evidence-based literacy assessment practices in contemporary educational settings.

The practical implications of the study lie in its potential to enhance literacy assessment practices. By providing a nuanced understanding of the strengths and weaknesses of various measurement models, educators and practitioners can make informed decisions about the most suitable approaches for their specific educational contexts. The study's emphasis on culturally responsive literacy assessment carries significant practical importance. Recognizing and adapting to cultural nuances in assessment strategies can lead to more contextually relevant and fair evaluations, ensuring that literacy assessments cater to the diverse skills and backgrounds of students.

On a social level, the study contributes to the promotion of equitable educational practices. By recognizing the limitations of traditional models and advocating for inclusive assessment frameworks, the findings encourage the development of tools that are unbiased and better accommodate the diverse skills and backgrounds of students. The study aligns with social contributions by fostering inclusivity, creating a learning environment that supports a sense of belonging for students from various cultural backgrounds.

In academic terms, the study advances knowledge in the field of literacy measurement. It acts as a valuable resource for researchers and scholars, providing a synthesized overview of classical and modern approaches. This contribution is vital for academics seeking to deepen their understanding of the complexities surrounding literacy assessment methodologies. By building on existing literature, the study adds to the academic discourse surrounding literacy assessment, contributing to the broader body of knowledge in the field.

Overall, the findings have the potential to improve educational outcomes by guiding educators towards more effective literacy assessment strategies. The recognition of diverse literacy skills and cultural influences contributes to the promotion of diversity and inclusion within educational settings. The study's multi-faceted contributions collectively aim to enhance literacy assessment practices in a diverse and dynamic educational landscape.

#### 5.0. Recommendation

The findings of this study on diverse approaches to literacy measurement warrant several recommendations to guide educators, practitioners, and policymakers in enhancing literacy assessment practices. Firstly, there is a need for professional development initiatives that familiarize educators with the strengths and

limitations of various literacy measurement models. Workshops and training programs should be designed to empower educators to choose assessment approaches that align with the diverse needs of their students and the specific contexts in which they teach.

Closely tied to this is the recommendation for the integration of culturally responsive literacy assessment practices into teacher training and professional development. Emphasizing cultural competence in assessment will better equip educators to recognize and address the unique linguistic and cultural influences that shape students' literacy skills. By fostering an understanding of diverse cultural perspectives, educators can create more inclusive and equitable assessment environments.

Furthermore, institutions and educational policymakers should consider investing in research and development initiatives aimed at creating innovative, contextually relevant literacy assessments. This involves supporting research that explores new models, incorporates performance-based tasks, and embraces technology to capture the dynamic and evolving nature of literacy skills. Collaborative efforts between researchers, educators, and assessment experts can lead to the design of assessments that better reflect the authentic literacy demands of the 21st century.

To address resource challenges associated with certain assessment approaches, it is recommended that educational institutions explore collaborative partnerships with external organizations, including non-profits, governmental agencies, and private enterprises. Such collaborations can provide additional resources for the implementation of performance-based assessments and culturally responsive approaches, ensuring sustainable and effective literacy measurement practices.

Lastly, ongoing evaluation and refinement of literacy assessment practices should be institutionalized. Regular reviews of assessment tools, methodologies, and their impact on diverse student populations can inform continuous improvement. This iterative process should involve collecting feedback from educators, students, and relevant stakeholders to ensure that literacy assessments remain relevant, fair, and aligned with the evolving landscape of education.

#### 6.0. Conclusion

In conclusion, the findings of this study contribute valuable insights to the discourse on literacy measurement, emphasizing the diversity of approaches and their implications for educational practices. Classical Test Theory, while foundational, revealed limitations in addressing the multidimensional nature of literacy skills, prompting a critical examination of its relevance in contemporary, diverse educational settings. Item Response Theory, particularly the Rasch model, emerged as a sophisticated framework but faced scrutiny for its assumptions, highlighting the need for a nuanced understanding of its applicability across diverse populations. The literature review underscored the evolution of literacy assessment, with performance-based and culturally responsive approaches offering innovative perspectives. While acknowledging challenges, the study emphasized the imperative of inclusivity and cultural sensitivity for fostering equitable educational outcomes. Positioned within this rich tapestry, the study aims to inform evidence-based literacy assessment practices in contemporary educational settings.

The practical implications of the study lie in its potential to enhance literacy assessment practices. Educators can make informed decisions about suitable approaches for diverse contexts, and the emphasis on culturally responsive assessment contributes to fair evaluations. On a social level, the study promotes inclusive educational practices, recognizing the diverse skills and backgrounds of students. Academically, the study advances knowledge in literacy measurement, providing a synthesized overview for researchers and scholars. Recommendations stemming from the findings emphasize the need for professional development, integrating culturally responsive practices, and investing in innovative, contextually relevant literacy assessments. Collaborative partnerships and ongoing evaluation processes are recommended to address resource challenges and ensure the continued improvement of literacy measurement practices. In essence, the study's multifaceted contributions offer a comprehensive understanding of literacy measurement, paving the way for informed practices and continued advancements in the field. The recommendations serve as practical guidelines for educators, policymakers, and researchers seeking to enhance literacy assessment in diverse educational settings.

# 7.0. References

- Abdullah, Z., Hoque, K. E., Ramlan, N. H., & Shafee, S. (2019). Designing the structural model of TVET lecturers' professionalism and generic skills based on an empirical study in Malaysia. Sage Open, 9(3), 2158244019861456.
- Ahmad, M., Badusah, J., Mansor, A. Z., Abdul Karim, A., Khalid, F., Daud, M. Y., ... & Zulkefle, D. F. (2016). The Application of 21st Century ICT Literacy Model among Teacher Trainees. Turkish Online Journal of Educational Technology-TOJET, 15(3), 151-161.
- Ahmad, S. A., Yoke, S. K., Yunos, R. M., & Amin, J. M. (2019). Exploring lecturers' readiness for 21st century education in Malaysian higher learning institutions. European Journal of Teaching and Education, 1(1), 15-29.
- Akil, W. F. H. W. M., & Adnan, A. H. M. (2022). Promoting Digital Literacy in Higher Education: Case Study of a Medical Laboratory Program in Malaysia. International Journal of Academic Research in Progressive Education and Development, 11(4), 735-755.
- 5. Andrich, D. (1988). Rasch Models for Measurement. Newbury Park, CA: SAGE Publications.
- Annoni, P., & Charron, N. (2019). Measurement assessment in cross-country comparative analysis: Rasch modelling on a measure of institutional quality. Social Indicators Research, 141(1), 31-60.
- Arabbani, F. K., Mulyani, S., Mahardiani, L., & Ariani, S. R. D. (2019, December). Analysis the quality of instrument for measuring chemical literacy abilities of high school student using Rasch model. In AIP Conference Proceedings (Vol. 2194, No. 1). AIP Publishing.
- Ayub, N. B., Ahsan, M. H. B., & Azman, N. S. B. (2021). Factors Determining Blended Learning Receptiveness Among Malaysian Academicians. Labuan Bulletin of International Business and Finance (LBIBF), 19(1), 118-133.
- Baharuddin, M. F., Masrek, M. N., Shuhidan, S. M., Razali, M. H., & Rahman, M. S. (2021). Evaluating the Content Validity of Digital Literacy Instrument for School Teachers in Malaysia through Expert Judgement.

- International Journal of Emerging Technology and Advanced Engineering, 11(07), 71-78.
- Baharuldin, Z., Jamaluddin, S., & Shaharom, M. S. N. (2019). The role of school administrative support and primary school teachers' ICT literacy to integrate ICT into the classrooms in Pahang, Malaysia. International Online Journal of Educational Leadership, 3(1), 26-42.
- 11. Baker, F. B. (2001). The Basics of Item Response Theory. College Park, MD: ERIC Clearinghouse on Assessment and Evaluation.
- 12. Chang, H. H., & Ying, Z. (1996). Computerized Adaptive Testing with the Rasch Model: Optimizing Precision of Ability Estimates. Journal of Educational Measurement, 33(3), 373-392.
- Chinedu, C. C., Saleem, A., & Wan Muda, W. H. N. (2023). Teaching and Learning Approaches: Curriculum Framework for Sustainability Literacy for Technical and Vocational Teacher Training Programmes in Malaysia. Sustainability, 15(3), 2543.
- 14. Crocker, L., & Algina, J. (1986). Introduction to classical and modern test theory. Rinehart & Winston.
- Darlington, R. B. (1975). Multiple regression in psychological research and practice. Psychological Bulletin, 82(2), 111–130.
- 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 (Indonesian Journal of Science Education), 9(3), 329-345.
- 17. De Clercq, B. (2019). A comparative analysis of the OECD/INFE financial knowledge assessment using the Rasch model. Empirical Research in Vocational Education and Training, 11(1), 1-29.
- Embretson, S. E., & Reise, S. P. (2000). Item Response Theory for Psychologists. Mahwah, NJ: Lawrence Erlbaum Associates.
- Farihah, U., Muhith, A., Saihan, S., & Ghani, F. A. (2020). Development of Reading Literacy Learning Model for Elementary School Students in Indonesia and Malaysia.
- 20. Fischer, L., Gnambs, T., Rohm, T., & Carstensen, C. H. (2019). Longitudinal linking of Rasch-model-scaled competence tests in large-scale assessments: A comparison and evaluation of different linking methods and anchoring designs based on two tests on mathematical competence administered in grades 5 and 7. Psychological Test and Assessment Modeling, 61(1), 37-64.
- Fountas, I. C., & Pinnell, G. S. (1996). Guided Reading: Good First Teaching for All Children. Portsmouth, NH: Heinemann.
- Freire, P., & Macedo, D. (1987). Literacy: Reading the Word and the World. South Hadley, MA: Bergin & Garvey.
- 23. Fuchs, D., & Fuchs, L. S. (2005). Peer-Assisted Learning Strategies: Promoting Word Recognition, Comprehension, and Fluency. Journal of Special Education, 39(1), 34–44.
- 24. Gardner, H. (1983). Frames of Mind: The Theory of Multiple Intelligences. New York: Basic Books.

- Gay, G. (2010). Culturally responsive teaching: Theory, research, and practice. Teachers College Press.
- Gie, T. A., & Fenn, C. J. (2019). Technology acceptance model and digital literacy of first-year students in a private institution of higher learning in Malaysia. BERJAYA Journal of Services & Management, 11, 103-116.
- 27. Goodman, K. (1986). What's Whole in Whole Language. Portsmouth, NH: Heinemann.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). Fundamentals of item response theory. Sage Publications.
- Hoi, V. N. (2020). Understanding higher education learners' acceptance and use of mobile devices for language learning: A Rasch-based path modeling approach. Computers & Education, 146, 103761.
- Howard, B. C., & Howard, I. P. (1982).
   Multidimensional Analysis of Abilities and Attitudes.
   Educational and Psychological Measurement, 42(4), 851–866.
- Islam, A. A., Gu, X., Crook, C., & Spector, J. M. (2020).
   Assessment of ICT in tertiary education applying structural equation modeling and Rasch model. Sage Open, 10(4), 2158244020975409.
- 32. Isnani, I., Utami, W. B., Susongko, P., & Lestiani, H. T. (2019). Estimation of college students' ability on real analysis course using Rasch model. Research and evaluation in education, 5(2), 95-102.
- 33. Isnani, I., Utami, W. B., Susongko, P., & Lestiani, H. T. (2019). Estimation of college students' ability on real analysis course using Rasch model. Research and evaluation in education, 5(2), 95-102.
- Jamaludin, K. A., Alias, N., DeWitt, D., & Ibrahim, M. M. (2020). Technical communication pedagogical model (TCPM) for Malaysian vocational colleges. Humanities and Social Sciences Communications, 7(1), 1-13.
- 35. Jamil, M. R. M., Hasyim, A. T. M., Othman, M. S., Ahmad, A. M., Noh, N. R. M., & Kamal, M. F. M. (2023). Digital Pedagogy Policy in Technical and Vocational Education and Training (TVET) in Malaysia: Fuzzy Delphi Approach. Journal of Technical Education and Training, 15(2), 1-10.
- Johnson, A. M., Robinson, M., Walton, D. M., & MacDermid, J. C. (2019). A comparison of the polytomous Rasch analysis output of RUMM2030 and R (ltm/eRm/TAM/lordif). BMC medical research methodology, 19, 1-12.
- 37. Ladson-Billings, G. (1995). But that's just good teaching! The case for culturally relevant pedagogy. Theory into Practice, 34(3), 159-165.
- 38. Ladson-Billings, G. (1995). But that's just good teaching! The case for culturally relevant pedagogy. Theory into Practice, 34(3), 159-165.
- Linacre, J. M. (1994). Sample Size and Item Calibration Stability. Rasch Measurement Transactions, 7(4), 328.
- 40. Linacre, J. M. (2002). Constructing a Rasch Rating Scale. Journal of Applied Measurement, 3(2), 151-162.
- 41. Linacre, J. M. (2002). How to Misinterpret DIF. Rasch Measurement Transactions, 15(4), 850.
- 42. Linacre, J. M. (2002). Invariance and Traceability. Rasch Measurement Transactions, 16(2), 877.

- 43. Linacre, J. M. (2002). What do Infit and Outfit, Mean-Square and Standardized Mean? Rasch Measurement Transactions, 16(2), 878.
- 44. Mahtari, S., Misbah, M., & Suryati, S. (2019). Analysis of the ability of High school students in solving science literacy questions based on the rasch model. Kasuari: Physics Education Journal (KPEJ), 2(1), 11-16.
- 45. Malaysia, H. U. K. (2020). Esl Lecturers'digital Competence In Malaysian Higher Education Institutions. International Journal of Future Generation Communication and Networking, 13(2), 415-425.
- 46. Martínez, R. A., & Fernández, R. L. (2015). Digital literacy for teachers in Cuban Tourism and Hotel Management Schools. Some experiences on its development. RUSC. Universities and Knowledge Society Journal, 12(3), 3-16.
- 47. Nallasamy, R., & Khairani, A. Z. B. (2022). Development and Validation of Reading Comprehension Assessments by Using GDINA Model. Malaysian Journal of Social Sciences and Humanities (MJSSH), 7(2), e001278-e001278.
- 48. Nashir, I. M., Mustapha, R., & Yusoff, A. (2019). Rasch Model Validity for Robotics Learning Survey in Technical and Vocational Education System. HONAI, 2(1), 47-58.
- Paivio, A. (1986). Mental Representations: A Dual Coding Approach. Oxford, England: Oxford University Press.
- Palincsar, A. S., & Brown, A. L. (1984). Reciprocal Teaching of Comprehension-Fostering and Comprehension-Monitoring Activities. Cognition and Instruction, 1(2), 117-175.
- Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). Knowing what students know: The science and design of educational assessment. National Academies Press.
- Planinic, M., Boone, W. J., Susac, A., & Ivanjek, L. (2019). Rasch analysis in physics education research:
   Why measurement matters. Physical Review Physics Education Research, 15(2), 020111.
- 53. Rahman, R. A., Bahiyah, P. T. S., & Jamil, H. (2022) Assessing the Level of Industrial Competence Possessed Among Polytechnics'lecturers And Its Necessitates Within The Tvet Industry Sector
- 54. Rencz, F., Mitev, A. Z., Szabó, Á., Beretzky, Z., Poór, A. K., Holló, P., ... & Brodszky, V. (2021). A Rasch model analysis of two interpretations of 'not relevant' responses on the Dermatology Life Quality Index (DLQI). Quality of Life Research, 30, 2375-2386.
- Robinson, M., Johnson, A. M., Walton, D. M., & MacDermid, J. C. (2019). A comparison of the polytomous Rasch analysis output of RUMM2030 and R (ltm/eRm/TAM/lordif). BMC medical research methodology, 19, 1-12.
- 56. Rubaai, N., & Hashim, H. (2021). Using cloud-based tools for English as a second language reading activities among Malaysian polytechnic students. Arab World English Journal (AWEJ) Special Issue on CALL, (7).
- Sanmugam, S. T., Rajanthran, N., & Periasamy, N. (2022). The Impact of Malaysian Polytechnic Tvet Lecturers Research Engagement on Their Professional Teaching And Learning Practices. International Journal

- of Technical Vocational and Engineering Technology (iJTveT), 3(2), 28-33.
- Sihombing, R. U., Naga, D. S., & Rahayu, W. (2019). A
  Rasch Model Measurement Analysis on Indonesian
  Science Literacy Test: Smart Way to Improve the
  Learning Assessment. IJER-INDONESIAN JOURNAL
  OF EDUCATIONAL REVIEW, 6(2).
- Sihombing, R. U., Naga, D. S., & Rahayu, W. (2019). A
  Rasch model measurement analysis on science literacy
  test of indonesian students: smart way to improve the
  learning assessment. IJER-Indonesian Journal of
  Educational Review, 6(1).
- Susongko, P., Kusuma, M., Arfiani, Y., Samsudin, A., & Aminudin, A. (2020). Revising of the Integrating Scientific Literacy Skills Scale (ISLS) with Rasch Model Analysis. Journal for the Education of Gifted Young Scientists, 8(4), 1583-1602.
- Vygotsky, L. S. (1978). Mind in Society: The Development of Higher Psychological Processes. Cambridge, MA: Harvard University Press.
- 62. Wan, X. N. A., & Borhan, M. T. B. (2021). Best Practice Construct for First Degree Supervision in Institute of Teacher Education (ITE) Malaysia: A Fuzzy Delphi Method (FDM) Analysis. Review of International Geographical Education Online, 11(4).
- Wright, B. D., & Masters, G. N. (1982). Rating Scale Analysis. Chicago, IL: MESA Press.
- 64. Wright, B. D., & Stone, M. H. (1979). Best Test Design. Chicago, IL: MESA Press.
- 65. Wright, B. D., & Stone, M. H. (1979). Best test design: Rasch measurement. MESA Press.
- 66. Wright, B. D., & Stone, M. H. (1979). Measurement Error and Bias in the Assessment of Intelligence. Psychological Bulletin, 86(2), 420-438.
- 67. Wright, B. D., & Stone, M. H. (1999). Measurement Essentials. Wilmington, DE: Wide Range, Inc.
- 68. Yudha, R. P. (2023). Higher order thinking skills (HOTS) test instrument: Validity and reliability analysis with the rasch model. EduMa: Mathematics education learning and teaching, 12(1), 21-38.
- Yusof, I. J., Latif, A. A., Derasid, N. A. C., & Jani, M. D. M. (2019). Research literacy level of education postgraduate research students using rasch measurement model. International Journal of Recent Technology and Engineering (IJRTE), 8, 791-796.
- Zulkifli, N., Hamzah, M. I., & Razak, K. A. (2022).
   Fuzzy Delphi Techniques: Creative Teaching Model Design for Polytechnic Islamic Education Lecturers.