International Journal on Advanced Science Engineering Information Technology

Vol.14 (2024) No. 6
ISSN: 2088-5334

# Big Data Analytics Quality Factors in Enhancing Healthcare Organizational Performance: A Pilot Study with Rasch Model Analysis

Wan Mohd Haffiz Mohd Nasir [a], Yusmadi Yah Jusoh [a,*], Rusli Abdullah [a], Salfarina Abdullah [a]

[a] *Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, Serdang, Selangor, Malaysia*
*Corresponding author: [*]yusmadi@upm.edu.my*

***Abstract*—Big Data Analytics (BDA) plays a pivotal role in the digital transformation of healthcare, significantly boosting organizational performance within the sector. As healthcare organizations increasingly adopt BDA to leverage data-driven decision-making, understanding the factors contributing to BDA quality becomes imperative. Thus, this study has proposed and developed the BDA quality conceptual model, and a pilot study is part of the process of completing the conceptual model development. The instrument, which is the questionnaire that has been designed, needs to be tested for reliability. Therefore, the pilot study aims to evaluate and refine the instrument used to assess BDA practitioners' comprehension of the constructs and the reliability of the items. This study utilized the probabilistic approach of Item Response Theory (IRT), explicitly employing the Rasch Measurement Model analysis to enhance the accuracy of measurement instruments, assess respondents' performance, and ensure instrument reliability. The survey instrument comprised 11 constructs and 64 items, which were designed to measure all the constructs: reliability, accuracy, completeness, timeliness, format, accessibility, usability, maintainability, portability, user satisfaction, and healthcare organizational performance. Data were collected from 20 respondents and synthesized according to their responses to each questionnaire item. The analyses were performed using Rasch analysis software, specifically Winsteps. The results of the Rasch analysis included findings on the reliability of persons and items, the distribution map of person-item relationships, identification of misfitting items, and assessment of unidimensionality. Ten items were removed from the initial set of 64 due to misfit, leaving 54 items that effectively measured respondents' understanding of BDA quality in healthcare organizational performance. Thus, Rasch measurement model analysis has confirmed the instrument was well constructed, valid, and reliable for actual study.**

***Keywords*— Big data analytics quality; BDA quality factors; pilot study; Rasch measurement model; Rasch analysis.**

*Manuscript received 27 Dec. 2023; revised 24 Aug. 2024; accepted 8 Nov. 2024. Date of publication 31 Dec. 2024.*


## I. Introduction

Big Data Analytics (BDA) plays a crucial role within organizations by enabling comprehensive management of the entire BDA life cycle. This life cycle encompasses various stages, including data acquisition, processing, management, storage, analytics, visualization, and the deployment of BDA applications across different platforms such as web and mobile applications [1], [2]. The analytics generated through BDA encompass descriptive and advanced forms such as predictive, prescriptive, and forecasting analytics [3], [4]. Unlike traditional Business Intelligence (BI), where analysis typically occurs towards the conclusion of the process, BDA enables analysis to commence early in the data acquisition phase, marking a significant departure in approach.

Over the past decade, there has been a substantial increase in the volume of data generated and collected across various activities, matched by advancements in technology that enhance our capability to analyze and interpret this data. The application of BDA in healthcare exemplifies its potential to contribute significantly to societal welfare [5]. BDA in healthcare refers to electronic health datasets that are exceedingly large and complex, posing challenges for traditional software, hardware, and data management tools [1], [6], [7]. In the healthcare sector, BDA is employed to reduce costs, predict epidemics, detect and treat diseases, improve quality of life, and prevent avoidable deaths [1], [3], [5], [8], [9]. The evolution of BDA has notably enhanced organizational performance, particularly within healthcare contexts [10].

Data-driven decision-making is becoming pivotal in advancing treatment approaches amidst global population growth and longer life expectancies. Early identification of significant illnesses is increasingly emphasized, aiming for more straightforward and cost-effective treatments when detected early in a patient's life [11]. Healthcare organizations worldwide face pressures to cut costs, improve coordination and outcomes, achieve greater efficiency, and prioritize patient-centered care [12]. These challenges underscore the urgent need to enhance the role of BDA in healthcare. Previous statistics reveal that the healthcare sector grapples with entrenched inefficiencies, contributing to annual wastage exceeding US$2 trillion [12]. Factors such as ineffective data acquisition, inadequate data sharing, and suboptimal information utilization significantly contribute to these inefficiencies [12], [13].

Healthcare organizations have embraced BDA as a comprehensive solution in the digital transformation era, facilitated by the widespread adoption of electronic medical records, healthcare information systems, and smart devices [14]. The volume of data in healthcare is rapidly escalating. It is projected to increase significantly [7], bringing to light various challenges regarding BDA quality within healthcare organizations, such as concerns related to reliability and accuracy [15]. The quality of BDA plays a crucial role in shaping outcomes in healthcare organizations, emphasizing the importance of delivering patient-centric treatments [16]. BDA's profound impact spans across all facets of the healthcare ecosystem, a reality that predates the COVID-19 pandemic. However, current challenges are magnified amidst the global health crisis, underscoring the heightened necessity for speed, agility, resilience, and precision, with unparalleled stakes [17].

To address these challenges, healthcare organizations must prioritize BDA quality as a critical opportunity and strategic imperative for enhancing organizational performance, effectively saving lives, and improving patients' quality of life [4], [14]. To optimize organizational effectiveness, the factors influencing BDA quality should serve as essential guidelines and mechanisms for assessing capabilities and performance characteristics, including maintainability and portability [18].

Given the urgency of healthcare crises, organizations must develop methods to evaluate BDA quality for care delivery. Understanding these factors enhances healthcare organizations' ability to improve healthcare organizational performance by ensuring care quality, reducing waste and errors, managing costs [19]–[21], enhancing healthcare facilities, and improving clinical outcomes for patients [19].

The comprehensive quality of BDA is crucial within its ecosystem, spanning acquisition, requirements, development, use, evaluation, and support. It significantly benefits developers and end users, establishing measures to ensure organizational performance success, particularly in healthcare [22], [23]. Despite the evolution of BDA implementation success, ongoing discussions persist regarding the factors influencing BDA quality. Numerous conceptual models have emerged from prior research, focusing on how these quality factors impact organizational performance from a firm's perspective [24]–[26].

Current models of BDA quality predominantly emphasize data and information quality, often overlooking holistic BDA quality aspects [27]. This limitation arises from a historical reliance on Information System (IS) theoretical models, neglecting the incorporation of Software Engineering (SE) theoretical frameworks [28]. BDA implementations typically serve as end-to-end solutions within their ecosystems, encompassing various related systems and applications. However, many of these BDA-related systems and applications frequently fall short of quality standards, fail to meet requirements, and struggle to satisfy user expectations, thus hindering organizational performance improvements [29], [30].

These BDA systems and applications commonly encounter scalability limitations, incomplete functionality, accuracy issues, maintenance complexities, deployment difficulties across different platforms, and comprehension challenges [31], [32]. Given BDA's role as an end-to-end solution within a complex ecosystem, existing theoretical models for studying BDA quality may not comprehensively capture all relevant quality factors [24].

This study devised an operational research framework structured around five sequential phases delineating research activities and anticipated outcomes aimed at achieving the research objectives. The first phase entailed a comprehensive literature review, and the second phase involved model development. The third phase involved conducting an empirical study, while the fourth phase focused on prototype development and model validation. The final phase encompassed compiling the research findings into a comprehensive report.

Before this pilot study, this research completed the Systematic Literature Review (SLR) in the first phase [33]. A conceptual BDA quality model has been proposed in the second phase with nine quality factors as determinants: reliability, accuracy, completeness, timeliness, format, accessibility, usability, maintainability, and portability. The quality factors have been defined, followed by the development of items to represent the quality factors. An expert review to verify the constructs and instruments has been done, and this will be followed by the pilot study, which will be explained further in this paper.

## II. Material and Method

This study employs a quantitative approach utilizing questionnaires designed to assess instrument reliability through a pilot study, following Pratt's recommendations [34]. A pilot study typically involves an initial, concise investigation using a small, convenient sample to refine questionnaires further [35], [36]. Reference [37] characterizes a pilot study as a scaled-down version of a full-scale study to increase the likelihood of success in subsequent research endeavors. Thus, the primary aim of this pilot study is to evaluate and improve the instrument's ability to measure BDA practitioners' conceptual comprehension of constructs and enhance item reliability.

"In validating the conceptual model of BDA quality, a quantitative approach was employed in this initial phase to assess the reliability of the study's instrument. The selection of quality factors was guided by their significance and prioritization in related studies of BDA and information systems and in software engineering. A key consideration throughout was how to enhance the survey instrument's quality. Reliability testing is crucial in survey-based research as it allows for statistical analysis, typically employing Cronbach's alpha to measure variables [38]. According to [39], instrument reliability refers to the consistency or stability of each measured item's observed score, comprising both the true and error score components."

The survey was conducted using Google Forms to collect data electronically. The distribution of the survey link was overseen by a liaison officer within selected organizations directly engaged in the BDA life cycle. Participants were drawn from public healthcare institutions in the Klang Valley, Selangor, Malaysia, adhering to predetermined selection criteria. Responses were measured on a five-point Likert scale, ranging from 1 ('strongly disagree') to 5 ('strongly agree'), to gauge agreement levels regarding the polytomous items included in the questionnaire [40].

The sampling approach for this study was adapted from [41], which recommended a sample size ranging from 10 to 40 participants. In this pilot study, 20 respondents completed the online questionnaire. Demographic variables of the pilot study participants were controlled, including gender, age, educational attainment, job title, and their experience or familiarity with various aspects of BDA systems and applications, such as data storage, processing, analytics, and visualization (e.g., MySejahtera app or other healthcare-related BDA systems/apps, COVID-19 or similar analytics and visualization tools).

TABLE I
DEMOGRAPHIC PROFILES OF THE RESPONDENTS IN PILOT STUDY (N=20)

| Measures | Category | Frequency | Percentage (%) |
|---|---|---|---|
| Gender | Male | 9 | 45 |
| | Female | 11 | 55 |
| Age (years) | 21 – 30 | 3 | 20 |
| | 31 - 40 | 8 | 45 |
| | 41 – 50 | 7 | 25 |
| | 51 - 60 | 2 | 10 |
| Involvement (Years) | 6 months to 3 | 10 | 50 |
| | 3 - 6 | 8 | 40 |
| | More than 6 | 2 | 10 |

Table I presents the demographic characteristics of the respondents. Nine were male (45%), and eleven were female (55%). The largest age group represented was 31 to 40 years old, including eight respondents (45%). This was followed by seven respondents (25%) in the 41 to 50 age bracket, three respondents (20%) in the 21 to 30 age range, and two respondents in the 51 to 60 age range.

Based on the treemaps visualization in Fig 1, in terms of educational attainment, respondents were classified into four categories: Postgraduate degree, Ph.D. (10%), master's degree (50%), Undergraduate degree, Bachelor's degree (25%), and Diploma (3%). The majority of respondents held postgraduate degrees. Regarding experience in the BDA life cycle, 50% of respondents had between 6 months to 3 years of experience, 40% had between 3 to 6 years of experience, and 10% had more than 6 years of experience. These demographic insights underscore participants' varied backgrounds and expertise levels in BDA applications within healthcare settings.

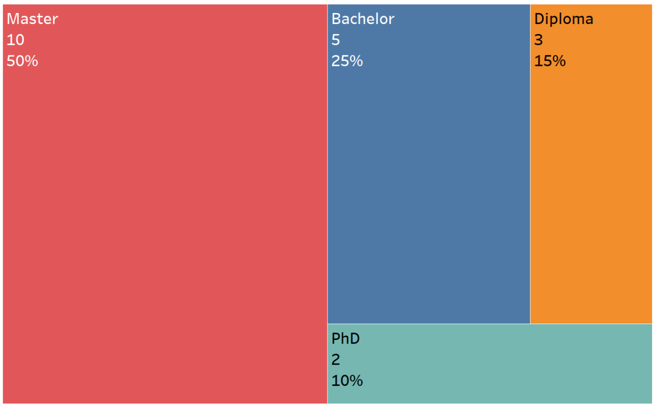


Fig. 1 Tree maps of Respondents Education Levels in Pilot Study (N=20)

Fig 2 visualizes the demographic characteristics of the respondents in terms of job roles. Respondents' roles were categorized into 10 categories: Certified Data Scientist (Statistician) (25%) with five respondents, Statistician (15%) with three respondents, two respondents (10%) for each Certified Data Scientist (IT Officer), Research Officer, IT Officer, and Assistant IT Officer. Followed by one respondent (5%) for each Specialist, Medical Officer, Senior Assistant Statistics Office, and IT Developer. This distribution illustrates the diverse roles in utilizing BDA solutions within healthcare organizations.

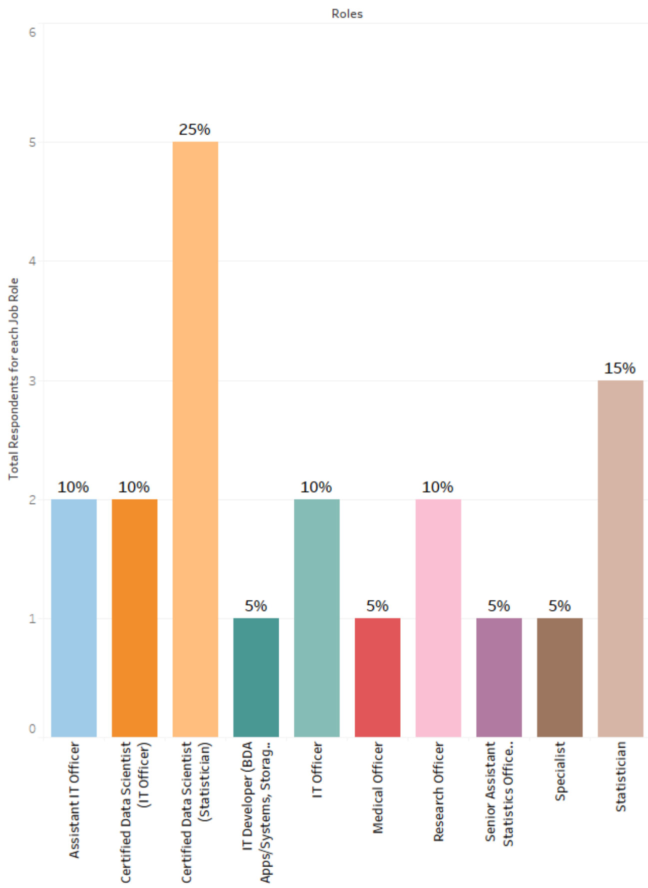


Fig. 2 Respondents Job Roles in Pilot Study (N=20)

In the pilot study, the Rasch Measurement Model (RMM) was employed to assess the reliability and item fit of the questionnaires used for data collection. The survey instrument consisted of 11 constructs and 64 items designed to evaluate various factors influencing Big Data Analytics (BDA) quality. These factors included reliability (5 items: R1 – R5), accuracy (5 items: A1 – A5), completeness (4 items: C1 – C4), timeliness (7 items: T1 – T7), format (4 items: F1 – F4), accessibility (4 items: Y1 – Y4), usability (4 items: U1 – U4), maintainability (6 items: M1 – M6), portability (6 items: P1 – P6), user satisfaction as a mediator (8 items: S1 – S8), and healthcare organizational performance (9 items: H1 – H9). This comprehensive instrument was utilized to gauge perceptions and assessments related to BDA quality across a range of dimensions within healthcare organizational contexts.

## III. Results and Discussion

Data from all 20 respondents were collected and aggregated to summarize their opinions on each item within the questionnaires. Analysis was performed using Rasch analysis software, precisely Winsteps 3.71.0.1 [42]. This study adopted the probabilistic approach of Item Response Theory (IRT) known as the Rasch Measurement Model [43], which is utilized to evaluate the reliability, validity, and appropriateness of scale measurement [44].

The Rasch Measurement Model employed in this pilot study assesses the reliability and validity of items and determines the appropriate scaling for measurement. It involves examining the ability of respondents (referred to as persons) to answer questionnaire items and simultaneously evaluating the difficulty levels of these items.

This methodology was typically employed to enhance the accuracy of construct instruments, assess respondents' performance, and monitor instrument quality. The subsequent section detailed the outcomes of Rasch analysis, encompassing computed metrics such as person and item reliability, person-item distribution mapping, item misfit assessment, and evaluation of unidimensionality [44].

### *A. Person and Item Reliability*

Fig. 3 and Fig. 4 present the results of person reliability and item reliability, respectively, indicating the instrument's internal consistency. Fig. 3 summarizes the person reliability among 20 respondents, yielding a score of 0.96, denoting 'excellent' reliability. The Cronbach's Alpha (KR-20) value of 0.97 further confirms 'excellent' internal consistency, consistent with findings from [45] and [46], which underscore the instrument's reliability in measuring BDA quality constructs.

| | Total Score | Count | Measure | Model Error | Infit | | Outfit | |
|---|---|---|---|---|---|---|---|---|
| | | | | | MNSQ | ZSTD | MNSQ | ZSTD |
| Mean | 273.2 | 64.0 | 1.13 | 2.48 | 1.05 | -0.3 | 1.00 | -0.6 |
| Standard Deviation | 25.6 | 0.0 | 0.99 | 1.40 | 0.64 | 3.4 | 0.6 | 3.0 |
| **Max** | 308.0 | 64.0 | 3.00 | 4.68 | 2.56 | 7.0 | 2.53 | 6.8 |
| **Min** | 216.0 | 64.0 | -0.38 | -0.41 | 0.23 | -6.3 | 0.23 | -6.4 |
| Real RMSE | 0.28 | True SD | 1.37 | Separation | 4.88 | Person Reliability | | 0.96 |
| Model RMSE | 0.25 | True SD | 1.38 | Separation | 5.59 | Person Reliability | | 0.97 |
| **S.E of Person Mean = 0.32** | | | | | | | | |
| Person Raw Score-To-Measure Correlation = 1.00 | | | | | | | | |
| **Cronbach Alpha (KR-20) Person Raw Score "Test" Reliability = 0.97** | | | | | | | | |

Fig. 3 Person Reliability of 20 Respondents

The Outfit Mean Square (MNSQ) value of 1.00 aligns perfectly with the expected value of 1, indicating a good fit. Similarly, the Outfit Standardized Mean Square (ZTSD) value of -0.6 approaches the expected value of 0, suggesting satisfactory adherence to Rasch model requirements. Person reliability assesses the varying abilities of individuals to score items accurately [44], [47]. Item reliability pertains to the consistency of item difficulty across different samples of the same size [44].

| | Total Score | Count | Measure | Model Error | Infit | | Outfit | |
|---|---|---|---|---|---|---|---|---|
| | | | | | MNSQ | ZSTD | MNSQ | ZSTD |
| Mean | 85.4 | 20.0 | 0.0 | 0.43 | 0.99 | 0.0 | 1.00 | 0.0 |
| Standard Deviation | 4.1 | 0.0 | 0.79 | 0.04 | 0.33 | 1.1 | 0.39 | 1.2 |
| **Max** | 96.0 | 20.0 | 1.43 | 0.62 | 1.63 | 1.9 | 2.00 | 2.8 |
| **Min** | 77.0 | 20.0 | -2.51 | 0.40 | 0.36 | -2.3 | 0.30 | -2.1 |
| Real RMSE | 0.46 | True SD | 0.65 | Separation | 1.40 | Item Reliability | | 0.66 |
| Model RMSE | 0.43 | True SD | 0.67 | Separation | 1.53 | Item Reliability | | 0.70 |
| **S.E of Item Mean = 0.10** | | | | | | | | |

Fig. 4 Item Reliability of 64 Items

Fig. 4, on the other hand, provides an overview of item reliability for 64 items in the instrument. The item reliability score of 0.66 indicates 'good' internal consistency, as per [46]. The Outfit Mean Square (MNSQ) value of 1.00 aligns precisely with the expected value of 1, indicating excellent fit. Similarly, the Outfit Standardized Mean Square (ZTSD) value of 0 aligns perfectly with the expected value of 0, further confirming strong adherence to the Rasch model. These results indicate that the items align well with the Rasch measurement requirements.

### *B. Person-Item Distribution Map*

This study employed a Rasch person-item distribution map (PIDM), the Wright Map, to assess the instrument's strengths and weaknesses. This map visually represents the distribution of item difficulty across a logarithmic scale, adjusted for respondent abilities (see Fig. 4). According to [44], the PIDM illustrates item difficulty from the easiest to the most challenging, providing insights into the alignment between respondent abilities and item difficulty levels. The mean person measure of 0.32 (S.E. of Person Mean) observed in Fig. 3 and 4 indicates that respondents generally favored endorsing higher item importance. This contrasts with the constrained Item Mean of 0.10 (S.E. of Item Mean), reflecting a discrepancy in respondent perceptions towards item difficulty levels.

As shown in Fig. 5, persons P12 and P16, identified as having the highest positions on the PIDM, consistently rated most items as highly important, indicating a tendency to endorse higher ratings. Conversely, Person P15 exhibited a lower ability to endorse items, consistently rating them lower across the scale. The range of respondent abilities fell between MAXPERSON = +4.68 logits and MINPERSON = -0.41 logits. The person separation index of 4.88 (see Fig. 3) suggests the presence of up to five distinct proficiency profiles, ranging from very high competence to very low competence.

Items, on the other hand, were distributed across a range from MAXITEM = +1.43 logits to MINITEM = -2.51 logits, with an item separation index of 1.40 (see Fig. 4). This indicates the instrument consists of items with two levels of difficulty: slightly difficult to moderate and easy items. The items were observed to cluster closely together, suggesting a scale size of 3.94 logits can effectively capture the range from -2.51 logits to +1.43 logits.

Among the items, S2 was identified as the most moderately difficult item, while item A5 was deemed the easiest to endorse. These findings provide a comprehensive overview of both respondent abilities and item difficulties within the studied instrument.

```
PERSON - MAP - ITEM
        <more>|<rare>
    5         +
              |
    P12 P16   |
              |
              |
    4   P5 P6 +
             S|
         P9   |
     P10 P7   |
              |
    3    P13  +
         P2   X  |
 P1 P20 P3 P4 |
             M|
              |
    2         +
              |
         P19  |T
         P11  |  S2
              |  M1      P5      P6      S5
    1        S+  M2      S6
  P14 P17 P18 |S C1      M5      S3      S4      U2
          P8  |  M3      P2      P4      U3
              |  C2      F1      F2      M4      M6      S8      U4
              |   A1       C3       C4       H1       H3       H5       P1       P3
S1
                 S7      U5      Y1
    0         +M F3      T5      T6
              |  A2      A3      A4      H2      H4      U1      Y2
      P15 T|  R2      T7      Y3      Y4
              |  H8      H9      R1      R4
              |S H7      T4
   -1         +  F4      H6      R3
              |
              |  F5      T1      T2      T3
              |T
              |  R5
   -2         +
              |
              |
              |  A5
              |
   -3         +
        <less>|<frequent>
```

Fig. 5 Wright Map for Person & Item

## C. Item Fit

In the subsequent Rasch analysis phase, the focus shifted towards identifying misfit items within the instrument. According to [46], item quality is deemed acceptable when the infit and outfit mean square (MnSq) values fall between 0.5 and 1.5. This study assessed misfit items for each construct, specifically examining those outside the acceptable range.

From the analysis presented in Fig. 6, based on the interpretation of infit mean square (MnSq) in the second row, items R2, A5, F5, U1, M2, M4, M6, P2, S2, and S5 were identified with infit MnSq values exceeding the acceptable range criteria [44]. Therefore, these 10 items were deemed misfits, and the recommendation was made to exclude them from the instrument. This approach aims to ensure the reliability and validity of the instrument by removing items that do not align with expected measurement standards.

| Constructs with the Item Codes | INFIT | | OUTFIT | | Pt Measure Corr. | Interpretation of INFIT MNSQ |
|---|---|---|---|---|---|---|
| | MNSQ | ZSTD | MNSQ | ZSTD | | |
| **Reliability (R)** | | | | | | |
| **R1** | 0.67 | -0.9 | 0.51 | -0.8 | 0.88 | Good |
| **R2** | 1.80 | 1.6 | 2.12 | 1.7 | 0.77 | **Eliminate** |
| **R3** | 1.10 | 0.4 | 0.87 | 0.0 | 0.74 | Good |
| **R4** | 0.83 | -0.3 | 0.94 | 0.1 | 0.85 | Good |
| **R5** | 1.12 | 0.4 | 0.98 | 0.4 | 0.83 | Good |
| **Accuracy (A)** | | | | | | |
| **A1** | 0.50 | -0.06 | 0.15 | -0.5 | 0.75 | Good |
| **A2** | 1.07 | 0.3 | 0.49 | 0.0 | 0.80 | Good |
| **A3** | 0.68 | -1.1 | 0.31 | -0.2 | 0.82 | Good |
| **A4** | 1.40 | 1.3 | 0.85 | 0.4 | 0.88 | Good |
| **A5** | 0.03 | -1.0 | 0.01 | -1.1 | 0.86 | **Eliminate** |
| **Completeness (C)** | | | | | | |
| **C1** | 0.87 | -0.2 | 0.44 | 0.0 | 0.94 | Good |
| **C2** | 0.99 | 0.1 | 0.55 | 0.1 | 0.91 | Good |
| **C3** | 0.98 | 0.1 | 0.50 | 0.0 | 0.95 | Good |
| **C4** | 0.97 | 0.0 | 0.49 | 0.0 | 0.91 | Good |
| **Timeliness (T)** | | | | | | |
| **T1** | 0.93 | 0.0 | 1.05 | 0.3 | 0.81 | Good |
| **T2** | 0.72 | -0.5 | 0.56 | -0.4 | 0.82 | Good |
| **T3** | 1.47 | 1.0 | 1.46 | 0.8 | 0.70 | Good |
| **T4** | 1.19 | 0.6 | 1.23 | 0.6 | 0.80 | Good |
| **T5** | 0.49 | -1.8 | 0.41 | -1.5 | 0.90 | Good |
| **T6** | 1.09 | 0.4 | 1.31 | 0.8 | 0.80 | Good |
| **T7** | 1.04 | 0.2 | 0.95 | 0.1 | 0.81 | Good |
| **Format (F)** | | | | | | |
| **F1** | 0.68 | -1.1 | 0.60 | -1.2 | 0.90 | Good |
| **F2** | 0.54 | -1.7 | 0.48 | -1.8 | 0.88 | Good |
| **F3** | 1.14 | 0.6 | 1.11 | 0.4 | 0.75 | Good |
| **F4** | 1.06 | 0.3 | 1.17 | 0.5 | 0.75 | Good |
| **F5** | 1.62 | 1.4 | 1.90 | 1.4 | 0.55 | **Eliminate** |
| **Accessibility (Y)** | | | | | | |
| **Y1** | 1.42 | 1.3 | 1.28 | 0.6 | 0.80 | Good |
| **Y2** | 1.15 | 0.5 | 0.78 | 0.3 | 0.87 | Good |
| **Y3** | 0.57 | -1.3 | 0.33 | -0.2 | 0.87 | Good |
| **Y4** | 0.58 | -1.2 | 0.33 | -0.2 | 0.93 | Good |
| **Usability (U)** | | | | | | |
| **U1** | 1.71 | 1.6 | 1.32 | 0.7 | 0.76 | **Eliminate** |
| **U2** | 0.86 | -0.2 | 0.86 | -0.1 | 0.92 | Good |
| **U3** | 0.58 | -0.9 | 0.42 | -1.1 | 0.96 | Good |
| **U4** | 0.80 | -0.2 | 0.68 | -0.4 | 0.93 | Good |
| **U5** | 0.75 | -0.3 | 0.95 | 0.1 | 0.92 | Good |
| **Maintainability (M)** | | | | | | |
| **M1** | 0.52 | -1.3 | 0.35 | -1.5 | 0.92 | Good |
| **M2** | 1.66 | 1.4 | 1.72 | 1.2 | 0.87 | **Eliminate** |
| **M3** | 0.03 | -1.0 | 0.1 | -1.2 | 0.95 | Good |
| **M4** | 1.82 | 1.6 | 2.18 | 1.8 | 0.68 | **Eliminate** |
| **M5** | 0.50 | -1.1 | 0.39 | -1.1 | 0.93 | Good |
| **M6** | 2.10 | 1.4 | 3.11 | 1.8 | 0.86 | **Eliminate** |
| **Portability (P)** | | | | | | |
| **P1** | 1.22 | .7 | 1.32 | .8 | 0.80 | Good |
| **P2** | 1.53 | 1.3 | 1.66 | 1.3 | 0.80 | **Eliminate** |
| **P3** | .74 | -0.7 | 0.62 | -.8 | 0.88 | Good |
| **P4** | .83 | -0.3 | 0.65 | -.6 | 0.91 | Good |
| **P5** | .63 | -0.9 | 0.50 | -1.0 | 0.92 | Good |
| **P6** | .98 | 0.1 | 1.01 | .2 | 0.86 | Good |
| **User Satisfaction (S)** | | | | | | |
| **S1** | 0.86 | -0.4 | 0.66 | -0.5 | 0.80 | Good |
| **S2** | 1.77 | 1.6 | 1.12 | 0.4 | 0.72 | **Eliminate** |
| **S3** | 0.36 | -1.9 | 0.24 | -1.7 | 0.90 | Good |
| **S4** | 0.96 | 0.0 | 0.94 | 0.1 | 0.83 | Good |
| **S5** | 1.69 | 1.2 | 2.01 | 1.3 | 0.83 | **Eliminate** |
| **S6** | 0.73 | -0.5 | 0.47 | -0.9 | 0.93 | Good |
| **S7** | 1.24 | 0.9 | 1.05 | 0.3 | 0.74 | Good |
| **S8** | 0.95 | 0.0 | 0.72 | -0.4 | 0.74 | Good |
| **Healthcare Organizational Performance (H)** | | | | | | |
| **H1** | 1.11 | 0.4 | 1.41 | 0.9 | 0.80 | Good |
| **H2** | 1.39 | 1.1 | 1.58 | 1.2 | 0.78 | Good |
| **H3** | 0.85 | -0.3 | 1.06 | 0.3 | 0.81 | Good |
| **H4** | 0.61 | -1.2 | 0.47 | -1.2 | 0.86 | Good |
| **H5** | 0.84 | -0.4 | 0.70 | -0.6 | 0.84 | Good |
| **H6** | 0.91 | -0.1 | 1.03 | 0.3 | 0.84 | Good |
| **H7** | 0.87 | -0.2 | 0.61 | -0.5 | 0.83 | Good |
| **H8** | 1.39 | 1.1 | 1.03 | 0.3 | 0.77 | Good |
| **H9** | 0.70 | -0.8 | 0.51 | -0.9 | 0.85 | Good |

Fig. 6 Item Fit

## D. Unidimensionality

Rasch analysis employs Principal Component Analysis (PCA) of residuals to assess the unidimensionality of the instrument. In the context of the Rasch model, unidimensionality refers to items within the questionnaires measuring a single construct [48]. Fig. 7 provides the results of the unidimensionality test for items, where 'Raw variance explained by measures' is reported at 42.6%. This value indicates that the instrument exhibits unidimensionality, as it exceeds the threshold of 40% variance explained [49]. The 'Unexplained variance in 1st contrast' value of 12% is considered 'good' according to [46], as it falls below the maximum acceptable threshold of 15%. This finding suggests that the instrument's items predominantly measure a single dimension, validating the construct's consistency [42]. Hence, the data align well with the Rasch model's estimation of measures, affirming the instrument's validity and reliability in measuring the intended construct.

| | | Empirical | | Modelled |
|---|---|---|---|---|
| Total raw variance in observations | 111.5 | 100.0% | | 100.0% |
| **Raw variance explained by measures** | 47.5 | **42.6%** | | 43.0% |
| Raw variance explained by persons | 27.1 | 24.3% | | 24.5% |
| Raw variance explained by items | 20.4 | 18.3% | | 18.4% |
| Raw unexplained variance (total) | 64.0 | 57.4% | 100.0% | 57.0% |
| **Unexplained variance in 1st contrast** | **13.4** | **12.0%** | 20.9% | |
| Unexplained variance in 2nd contrast | 7.6 | 6.8% | 11.9% | |
| Unexplained variance in 3rd contrast | 6.8 | 6.1% | 10.6% | |
| Unexplained variance in 4th contrast | 5.6 | 5.0% | 8.7% | |
| Unexplained variance in 5th contrast | 5.2 | 4.6% | 8.1% | |

Fig. 7 Standardized Residual Variance (in Eigenvalue Units)

*E. Scale Calibration*

According to [44], it is crucial to calibrate the rating scale during the pilot phase of instrument development to ensure its suitability and effectiveness in measuring the intended constructs. Reference [42] has outlined five criteria for diagnosing problematic rating scales, including 1) distinct peaks in the probability curve for each category, 2) increasing observed average values as categories ascend, and 3) a separation (S') of structure calibration (SC) differences between thresholds falling within the range of $1.0 < x < 5.0$. Failure to meet these criteria may indicate the need to collapse or combine rating scale categories [42].

Another essential assumption for optimizing the effectiveness of rating scale categories is the unimodal frequency distribution, which resembles a series of hills under the first criteria mentioned above—the probability curve displaying distinct peaks for each category, as illustrated in Fig. 8. This ensures that the rating scale effectively captures and distinguishes between different levels of response within the measured construct.

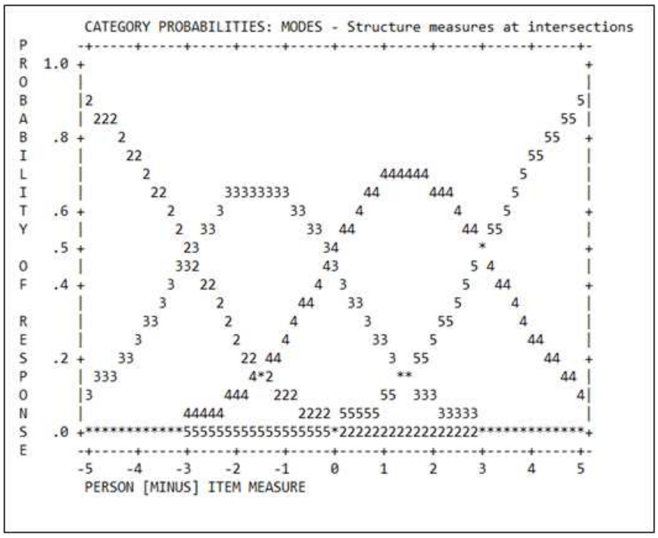


Fig. 8 Scale Probability Curve

As per [44], scale calibration during pilot testing offers empirical insights into respondents' comprehension of various scaling labels. In the questionnaires, the 5-point Likert scale was designated as Category 1, Category 2, Category 3, Category 4, and Category 5 within Rasch's scale rating system. This approach ensures clarity and consistency in evaluating respondents' perceptions across different levels of the scale, as validated through empirical testing.

To maintain the integrity of scale categories without the need for collapsing or combining, the separation (S') of structure calibration (SC) differences must fall within the range of 1.4 to 5 logits, as outlined in the third criterion. Fig. 9 displays the outcome of the scaling calibration process. The separation (S') value is computed as follows:

$$\text{Sn-(n+1)} = \text{SCCat(n)} - \text{SCCat(n+1)} \quad (1)$$

The separation (S) denotes the rating value that measures the difference between the structure calibration (SC) of one category (n) and the structure calibration of the subsequent category (n+1). This metric assesses how distinctively each category on the scale captures varying levels of response or agreement, ensuring the scale effectively discriminates between different degrees of perception or attitude.

| Category | | Observed | | Observed Average | Sample Expect | Infit | Outfit | Structure Calibration | Category Measure |
|---|---|---|---|---|---|---|---|---|---|
| Label | Score | Count | % | | | MNSQ | MNSQ | | |
| **1** | 1 | None | | | | | | | |
| 2 | 2 | 8 | 1 | 0.00 | -0.39 | 1.19 | 1.21 | None | (-4.06) |
| 3 | 3 | 155 | 12 | 0.81 | 0.56 | 1.35 | 1.56 | -2.92 | -1.51 |
| 4 | 4 | 602 | 47 | 1.90 | 2.05 | 0.81 | 0.70 | -0.09 | 1.47 |
| 5 | 5 | 515 | 40 | 3.69 | 3.59 | 0.89 | 3.01 | 3.01 | (4.14) |

Fig. 9 Scaling Categories

The criterion for maintaining rating scale categories without collapsing or combining them typically falls within the 1.4 to 5 logits range. After assessing scale ratings, calculations confirm that all categories can indeed be retained. This ensures that each category effectively contributes to distinguishing between different levels of responses within the measured construct, thereby supporting the validity and reliability of the rating scale.

Scale 2 & 3: The difference is 2.92 logits.

Scale 3 & 4: The difference is 2.83 logits.

Scale 4 & 5: The difference is 3.10 logits.

These calculations demonstrate the differences between adjacent scale categories in logits, indicating that each category is sufficiently distinct and supports the retention of all scale categories without the need for collapsing or combining them.

Based on the findings, the five-point Likert scale configuration will be preserved and utilized in the subsequent phases of the study. This scale categorizes responses as follows: 1 signifies 'Strongly Disagree', 2 denotes 'Disagree', 3 represents 'Neutral', 4 indicates 'Agree', and 5 signifies 'Strongly Agree'. This decision is supported by the validation of the scale's effectiveness in distinguishing varying degrees of agreement or disagreement among respondents, ensuring its suitability for measuring the intended constructs in the actual study.

*F. Rasch Analysis Summary for Pilot Study*

The Rasch Model Analysis was employed to analyze the data gathered from the pilot study. According to the results in Fig. 6, ten items (R2, A5, F5, U1, M2, M4, M6, P2, S2, S5) out of 64 were identified as misfitting items and subsequently removed, leaving 54 items that remained suitable. These retained items were deemed effective in assessing respondents' competencies, as they met all Rasch assumptions.

## IV. Conclusion

The study findings suggest a congruent alignment between the distribution of respondents and the items measuring their comprehension of BDA quality in healthcare organizational performance. Additionally, the high Cronbach's Alpha value indicates strong internal consistency among the scale items. Therefore, the Rasch measurement model analysis affirms that the instrument is well-constructed, valid, and reliable for use in subsequent studies.

## Acknowledgment

The authors thank the Faculty of Computer Science and Information Technology (FSKTM) and the School of Graduate Studies (SGS), Universiti Putra Malaysia (UPM), which supported this work.

## References

[1] D. Tosi, R. Kokaj, and M. Roccetti, "15 years of Big Data: a systematic literature review," *Journal of Big Data*, vol. 11, no. 1, May 2024, doi:10.1186/s40537-024-00914-9.

[2] I. Taleb, M. A. Serhani, and R. Dssouli, "Big Data Quality: A Survey," *2018 IEEE International Congress on Big Data (BigData Congress)*, pp. 166–173, Jul. 2018, doi: 10.1109/bigdatacongress.2018.00029.

[3] S. Sarker, M. S. Arefin, M. Kowsher, T. Bhuiyan, P. K. Dhar, and O.-J. Kwon, "A Comprehensive Review on Big Data for Industries: Challenges and Opportunities," *IEEE Access*, vol. 11, pp. 744–769, 2023, doi: 10.1109/access.2022.3232526.

[4] A. Rehman, S. Naz, and I. Razzak, "Leveraging big data analytics in healthcare enhancement: trends, challenges and opportunities," *Multimedia Systems*, vol. 28, no. 4, pp. 1339–1371, Jan. 2021, doi:10.1007/s00530-020-00736-8.

[5] A. Ahmed, R. Xi, M. Hou, S. A. Shah, and S. Hameed, "Harnessing Big Data Analytics for Healthcare: A Comprehensive Review of Frameworks, Implications, Applications, and Impacts," *IEEE Access*, vol. 11, pp. 112891–112928, 2023, doi: 10.1109/access.2023.3323574.

[6] P. Goyal and R. Malviya, "Challenges and opportunities of big data analytics in healthcare," *Health Care Science*, vol. 2, no. 5, pp. 328–338, Oct. 2023, doi: 10.1002/hcs2.66.

[7] W. Raghupathi and V. Raghupathi, "Big data analytics in healthcare: promise and potential," *Health Information Science and Systems*, vol. 2, no. 1, Feb. 2014, doi: 10.1186/2047-2501-2-3.

[8] P. B. Dharmawan, I. G. A. N. S. Maharani, and C. Tho, "Big Data Capabilities for Hospital: A Systematic Literature Review," *Procedia Computer Science*, vol. 227, pp. 272–281, 2023, doi:10.1016/j.procs.2023.10.525.

[9] K. E. Alhajaj and I. A. Moonesar, "The power of big data mining to improve the health care system in the United Arab Emirates," *Journal of Big Data*, vol. 10, no. 1, Feb. 2023, doi: 10.1186/s40537-022-00681-5.

[10] Y. Wang, L. Kung, S. Gupta, and S. Ozdemir, "Leveraging Big Data Analytics to Improve Quality of Care in Healthcare Organizations: A Configurational Perspective," *British Journal of Management*, vol. 30, no. 2, pp. 362–388, Apr. 2019, doi: 10.1111/1467-8551.12332.

[11] K. Batko and A. Ślęzak, "The use of Big Data Analytics in healthcare," *Journal of Big Data*, vol. 9, no. 1, Jan. 2022, doi: 10.1186/s40537-021-00553-4.

[12] J. W. Cortada, D. Gordon, and B. Lenihan, The Value of Analytics in Healthcare. IBM Institute for Business Value, Armonk, NY, USA, 2012, Report No.: GBE03476-USEN-00.

[13] N. Côrte-Real, P. Ruivo, and T. Oliveira, "Leveraging internet of things and big data analytics initiatives in European and American firms: Is data quality a way to extract business value?," *Information & Management*, vol. 57, no. 1, p. 103141, Jan. 2020, doi:10.1016/j.im.2019.01.003.

[14] N. Mehta and A. Pandit, "Concurrence of big data analytics and healthcare: A systematic review," *International Journal of Medical Informatics*, vol. 114, pp. 57–65, Jun. 2018, doi:10.1016/j.ijmedinf.2018.03.013.

[15] H. Kang and S. Sibbald, "Challenges to Using Big Data in Health Services Research," *University of Western Ontario Medical Journal*, vol. 87, no. 2, pp. 18–20, Mar. 2019, doi: 10.5206/uwomj.v87i2.1140.

[16] G. Manikandan, S. Abirami, K. Gokul, and G. Deepalakshmi, "Big data analytics in healthcare," *Big Data Analytics for Healthcare*, pp. 3–11, 2022, doi: 10.1016/b978-0-323-91907-4.00008-x.

[17] Cloudera, "Welcome to the Revolution in Healthcare," [Online]. Available: https://fr.cloudera.com/content/dam/www/marketing/resources/solution-briefs/cloudera-ibm-welcome-to-the-revolution-in-healthcare.pdf?daqp=true. [Accessed: Aug. 24, 2024].

[18] J. A. McCall, P. K. Richards, and G. F. Walters, *Factors in Software Quality, Concept and Definitions of Software Quality*, 1977. [Online]. Available: http://www.dtic.mil/dtic/tr/fulltext/u2/a049014.pdf. Accessed: Aug. 24, 2024.

[19] R. Raja, I. Mukherjee, and B. K. Sarkar, "A Systematic Review of Healthcare Big Data," *Scientific Programming*, vol. 2020, pp. 1–15, Jul. 2020, doi: 10.1155/2020/5471849.

[20] N. Mehta, A. Pandit, and M. Kulkarni, "Elements of Healthcare Big Data Analytics," *Big Data Analytics in Healthcare*, pp. 23–43, Oct. 2019, doi: 10.1007/978-3-030-31672-3_2.

[21] Y. Wang, "Leveraging Big Data Analytics to Improve Quality of Care In Health Care: A fsQCA Approach," *Proceedings of the 51st Hawaii International Conference on System Sciences*, 2018, doi:10.24251/hicss.2018.097.

[22] ISO, "ISO/IEC 25010:2011 (SQuaRE) Quality Model," Geneva, Switzerland, 2011. [Online]. Available: https://www.iso.org/obp/ui/#iso:std:iso-iec:25010:ed-1:v1:en.

[23] B. W. Boehm, J. R. Brown, H. Kaspar, M. Lipow, G. McLeod and M. Merritt, "Characteristics of Software Quality," North Holland, 1978.

[24] S. Fosso Wamba, S. Akter, and M. de Bourmont, "Quality dominant logic in big data analytics and firm performance," *Business Process Management Journal*, vol. 25, no. 3, pp. 512–532, Jun. 2018, doi:10.1108/bpmj-08-2017-0218.

[25] S. F. Wamba, A. Gunasekaran, S. Akter, S. J. Ren, R. Dubey, and S. J. Childe, "Big data analytics and firm performance: Effects of dynamic capabilities," *Journal of Business Research*, vol. 70, pp. 356–365, Jan. 2017, doi: 10.1016/j.jbusres.2016.08.009.

[26] S. Ji-fan Ren, S. Fosso Wamba, S. Akter, R. Dubey, and S. J. Childe, "Modelling quality dynamics, business value and firm performance in a big data analytics environment," *International Journal of Production Research*, vol. 55, no. 17, pp. 5011–5026, Mar. 2016, doi:10.1080/00207543.2016.1154209.

[27] P. H. Shariat Panahy, F. Sidi, L. S. Affendey, M. A. Jabar, H. Ibrahim, and A. Mustapha, "A Framework to Construct Data Quality Dimensions Relationships," *Indian Journal of Science and Technology*, vol. 6, no. 5, pp. 1–10, May 2013, doi:10.17485/ijst/2013/v6i5.10.

[28] M. Wook, "Big Data Analytics Application Model Based on Data Quality Dimensions and Big Data Traits in Public Sector," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 2, pp. 1247–1256, Apr. 2020, doi:10.30534/ijatcse/2020/53922020.

[29] A. Ameen, D. Al-Ali, O. Isaac, and F. Mohammed, "Examining relationship between service quality, user satisfaction, and performance impact in the context of smart government in UAE," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 6, p. 6026, Dec. 2020, doi:10.11591/ijece.v10i6.pp6026-6033.

[30] S. Fosso Wamba, S. Akter, L. Trinchera, and M. De Bourmont, "Turning information quality into firm performance in the big data economy," *Management Decision*, vol. 57, no. 8, pp. 1756–1783, Sep. 2019, doi: 10.1108/md-04-2018-0394.

[31] S. Akter, S. Fosso Wamba, and S. Dewan, "Why PLS-SEM is suitable for complex modelling? An empirical illustration in big data analytics quality," *Production Planning & Control*, vol. 28, no. 11–12, pp. 1011–1021, Jul. 2017, doi: 10.1080/09537287.2016.1267411.

[32] M. Ghasemaghaei and G. Calic, "Can big data improve firm decision quality? The role of data quality and data diagnosticity," *Decision Support Systems*, vol. 120, pp. 38–49, May 2019, doi: 10.1016/j.dss.2019.03.008.

[33] W. M. H. M. Nasir, Y. Y. Jusoh, R. Abdullah, and S. Abdullah, "Towards Healthcare Organizational Performance Deriving by Big Data Analytics Quality Factors: A Systematic Literature Review," *2022 Applied Informatics International Conference (AiIC)*, pp. 1–6, May 2022, doi: 10.1109/aiic54368.2022.9914026.

[34] D. Pratt, Curriculum Design and Development. Harcourt Brace, 1980.

[35] M.-C. Boudreau, D. Gefen, and D. W. Straub, "Validation in Information Systems Research: A State-of-the-Art Assessment," *MIS Quarterly*, vol. 25, no. 1, p. 1, Mar. 2001, doi: 10.2307/3250956.

[36] D. Straub and D. Gefen, “Validation Guidelines for IS Positivist Research,” *Communications of the Association for Information Systems*, vol. 13, 2004, doi: 10.17705/1cais.01324.
[37] E. R. Van Teijlingen, A. Rennie, V. Hundley, and W. Graham, “The importance of conducting and reporting pilot studies: the example of the Scottish Births Survey,” *Journal of Advanced Nursing*, vol. 34, no. 3, pp. 289–295, May 2001, doi: 10.1046/j.1365-2648.2001.01757.x.
[38] J. A. Gliem and R. R. Gliem, "Calculating, interpreting, and reporting Cronbach’s alpha reliability coefficient for Likert-type scales," in Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education, Ohio State University, Columbus, OH, USA, 2003.
[39] W. Wiersma and S. G. Jurs, Research Methods in Education: An Introduction, 9th edition. Pearson, 2009.
[40] D. Andrich, “A rating formulation for ordered response categories,” Psychometrika, vol. 43, no. 4, pp. 561–573, Dec. 1978, doi: 10.1007/bf02293814.
[41] M. A. Hertzog, “Considerations in determining sample size for pilot studies,” *Research in Nursing & Health*, vol. 31, no. 2, pp. 180–191, Jan. 2008, doi: 10.1002/nur.20247.
[42] J. M. Linacre, A User's Guide to Winsteps/Ministep: Rasch-Model Computer Programs. 2015. [Online]. Available: http://www.winsteps.com/manuals.htm.
[43] G. Rasch, Probabilistic Models for Some Intelligence and Attainment Tests. Copenhagen: Danish Institute for Educational Research, 1960.
[44] C. Bond, T. & Fox, Applying the Rasch Model: Fundamental Measurement in the Human Sciences. Mahwah, NJ: Lawrence Erlbaum Associates, 2007.
[45] I. H. Nunnally, J. C., & Bernstein, Psychometric Theory, 3rd edn. New York: New York, NY: McGraw-Hill, 1994.
[46] W. P. Fisher, “Rating Scale Instrument Quality Criteria,” Rasch Meas. Trans., vol. 21, no. 1, p. 1095, 2007.
[47] G. N. Wright, B.D. & Masters, Rating Scale Analysis: Rasch Measurement. Chicago: Mesa Press, 1982.
[48] E. Brentari, S. Golia, and M. Manisera, "Models for categorical data: A comparison between the Rasch model and nonlinear principal component analysis," *Statistica & Applicazioni*, vol. 5, no. 1, pp. 53–77, 2007.
[49] A. A. Aziz, M. S. Masodi and A. Zaharim, Asas Model Pengukuran Rasch: Pembentukan Skala Dan Struktur Pengukuran, Bangi, Malaysia, Penerbit Universiti Kebangsaan Malaysia, 2017.