**MUMTAZ**: Jurnal Pendidikan Agama Islam
Volume 5 Nomor 3 Tahun 2025
E-ISSN 2809-204X
P-ISSN 2809-2139

# MANY-FACET RASCH MODEL (MFRM) FOR DIGITAL CITIZENSHIP MEASUREMENT ANALYSIS IN ISLAMIC SCHOOL: A BIBLIOMETRIC REVIEW

Vita Fitriatul Ulya*[1], Fikiy Afriyanto[2], Suparto[3], Kusaeri[4], Vacharee Dedduang[5]
[1*] Islamic Religious Education Faculty, Universitas Al-Hikmah Indonesia, Tuban, Indonesia
[2,3,4] Islamic Religious Education Faculty, Universitas Islam Negeri Sunan Ampel, Surabaya, Indonesia
[5] School of Educational Studies, Sukhothai Thammatirath Open Univeristy, Nonthaburi, Thailand
e-mail: [1]vitaf3@gmail.com, [2]fikyafry@gmail.com, [3]suparto.zakfikam@uinsa.ac.id, [4]kusaeri@uinsa.ac.id, [5]sna_ded@hotmail.com

| Received: October 11, 2025 | Revised: November 24, 2025 | Accepted: December 17, 2025 | Published: December, 2025 |
|---|---|---|---|

*Corresponding author

## Abstract

*The increasing need for Digital Citizenship competencies in education requires the availability of assessment instruments that are objective, valid, and in line with the challenges of digital ethics in the modern era. This study aims to map the global development of the use of the Many-Facet Rasch Model (MFRM) in the analysis of Digital Citizenship (DC) instruments and identify the implications of its application for the development of assessments in Islamic schools. Using the Scopus and ERIC data-based bibliometric analysis design, this study processed 143 publications through the bibliometrix device at R-Studio, including publication trend analysis, scientific collaboration, thematic structure, and citation network. The results of the study show that the integration of MFRM in digital citizenship assessment provides a strong methodological foundation to overcome the issue of rater bias and strengthen the objectivity of performative evaluation, especially in the context of Islamic Schools which require digital character assessments based on moral values, morals, and ethical responsibility in accordance with Islamic education principles. This study concludes that MFRM is a strategic approach to improve the quality of DC assessments in a fair and value-based manner, as well as open up opportunities for innovation of digital character evaluation instruments in Islamic schools. Further studies are recommended to test the empirical implementation of MFRM-based instruments and Differential Facet Functioning (DFF) analysis in madrasah and cross-cultural contexts.*

**Keywords:** *Bibliometric Review; Digital Citizenship; Islamic School; Many-Facet Rasch Model; Measurement Analysis.*

**Abstrak**

*Meningkatnya kebutuhan akan kompetensi Digital Citizenship (kewargaan digital) dalam pendidikan menuntut tersedianya instrumen asesmen yang objektif, valid, serta sesuai dengan tantangan etika digital di era modern. Penelitian ini bertujuan untuk memetakan perkembangan global penggunaan Many-Facet Rasch Model (MFRM) dalam analisis instrumen Digital Citizenship (DC) serta mengidentifikasi implikasi penerapannya bagi pengembangan asesmen di sekolah Islam. Dengan menggunakan desain analisis bibliometrik berbasis data Scopus dan ERIC, penelitian ini mengolah 143 publikasi melalui perangkat bibliometrix di R-Studio, mencakup analisis tren publikasi, kolaborasi ilmiah, struktur tematik, dan jaringan sitasi. Hasil penelitian menunjukkan bahwa integrasi MFRM dalam asesmen kewargaan digital memberikan landasan metodologis yang kuat untuk mengatasi masalah bias penilai dan memperkuat objektivitas evaluasi performatif, khususnya dalam konteks sekolah Islam yang membutuhkan asesmen karakter digital berbasis nilai moral, akhlak, dan tanggung jawab etis sesuai dengan prinsip pendidikan Islam. Penelitian ini menyimpulkan bahwa MFRM merupakan pendekatan strategis untuk meningkatkan kualitas asesmen DC secara adil dan berbasis nilai, sekaligus membuka peluang inovasi instrumen evaluasi karakter digital di sekolah Islam. Studi lanjutan direkomendasikan untuk menguji implementasi empiris instrumen berbasis MFRM serta analisis Differential Facet Functioning (DFF) di madrasah dan konteks lintas budaya.*

**Kata Kunci:** *Bibliometric Review; Digital Citizenship; Islamic School; Many-Facet Rasch Model; Measurement Analysis.*

## A. Introduction

The massive increase in internet access and digital platforms globally has placed Digital Citizenship (DC*)* as a crucial pillar in the education system (Morales-álvarez et al., 2025; Naeem & Mushibwe, 2025; Webster, 2025). This concept *of DC* is not just digital literacy, but encompasses the ability of individuals to engage in digital society in a positive, responsible, ethical, and safe manner (Pangrazio & Sefton-Green, 2021) (Özüdog, 2025). Various countries and educational institutions have integrated *DC* as a core competency that every individual must master, given the ever-increasing risks such as *cyberbullying*, disinformation, and privacy issues (Iskandar et al., 2025; Reilly et al., 2024). Therefore, significant investments have been allocated to develop valid and reliable instruments to measure this level of competence, ensuring that educational interventions are based on accurate and reliable data (Chen et al., 2021).

Ideally, DC measurement instruments, which are often multidimensional and use *rating scales* or observation-based rubrics, should produce scores that truly

reflect the ability of respondents, free from the influence of *disruptive factors* (Choi et al., 2017; Dass & Kumar, 2025; Feriandi et al., 2022; Jones & Mitchell, 2015). The researchers hope that the instrument can guarantee the objectivity and fairness of the assessment. But in reality, there is a significant gap between this ideal expectation and existing practice(Valentine et al., 2022; Zlatkin-troitschanskaia et al., 2019). Measurement in the educational domain is very susceptible to *bias* and *variability* caused by external factors or so-called *facets*, such as the difference in rigidity between one rater (*rater*) and another (Arifiyanti et al., 2023; Goodwin, 2016) inconsistency in the use of rating scale categories (Dogan & Uluman, 2017), or even the impact of unmanaged cultural contexts (Veenhoven, 2012). This gap casts doubt on the validity of the findings, and ultimately, undermines the effectiveness of *well-designed* DC intervention programs.

To address *the vulnerability of DC* instruments to bias, psychometricians have offered several remedial approaches. To ensure item quality and unidimensionality, a One Parameter Rasch Model is used that is able to separate the difficulty of the item from the ability of the respondent(Heru et al., 2022; Laura-de-la-cruz et al., 2024). In the context of observation-based assessment, assessor bias is addressed through assessor training or *Generalizability Theory* analysis to estimate variance between assessors (Brennan, 2011; Clayson et al., 2021). Meanwhile, the inconsistency in the use of rating categories is addressed with special models such as *the Rating Scale Model (RSM)* or *Partial Credit Model (PCM)* (Yuanyuan et al., 2025). However, because each of these approaches works separately and only partially solves the problem, a framework is needed that is capable of handling items, respondents, assessors, and scale categories simultaneously.

From a series of such solution offerings, experts are now narrowing down to the use of *the Many-Facet Rasch Model (MFRM)* as the most comprehensive solution (Aslanoglu et al., 2020). The MFRM, developed by Linacre, is capable of integrating all relevant sources of variation (respondents, items, assessors, and scales) simultaneously in a single logistical equation (Bahrouni, 2016; John & Linacre, 1993). This model fundamentally allows researchers to obtain an estimate of respondents' abilities and item difficulty that is clean from the influence of appraiser bias and other disruptive facets(Ho, 2019; Mclaughlin et al., 2023). Therefore, the adoption of MFRM is not just a methodological choice, but rather an offer of the best solution to achieve maximum measurement objectivity and consequences-based validity in the analysis of *DC* instruments on the global stage. The *suitability of the Many-Facet Rasch Model (MFRM)* with *the* analysis of DC instruments is not only based on its statistical sophistication, but especially on the model's alignment with the inherent nature of digital competence itself (Sata & Karakaya, 2020). DC

measurements are almost always multidimensional and contextual, encompassing cognitive (understanding of rights and responsibilities), affective (ethical attitudes), and psychomotor (safe behavior practices) (Connolly & Miller, 2022).

The main reason the application of *the Many-Facet Rasch Model (MFRM)* is so appropriately used in *the analysis of DC instruments* is that the measurement characteristics in this field often involve performance-based assessments, such as the ability to interact ethically in the digital space or the practice of maintaining data security (Sebasti & Lozano-d, 2021). This kind of instrument is generally assessed by raters using rubrics, so it is vulnerable to individual subjectivity. MFRM is able to model and correct variability due to assessor bias, resulting in a more objective estimation of ability (Guler, 2014). In addition, the K *DC* instrument is widely used in different countries and cultures, so variations in digital contexts can act as a disruptive facet (Nordin et al., 2016). MFRM provides a mechanism to test systematic differences in item difficulty and assessment rigor between demographic groups or between countries through *Differential Facet Functioning (DFF) analysis*, thereby increasing the validity of cross-cultural interpretations (Effendi et al., 2020). Furthermore, with its ability to place respondent abilities, item difficulties, and evaluator characteristics on a single logit scale of intervals, MFRM ensures that the evaluation and data-driven decision-making process takes place fairly and objectively (Wolfe, 2015). Thus, MFRM is not only an alternative to statistical analysis, but it becomes an important methodological approach to ensure accuracy, objectivity, and fairness in *the assessment of DCs*, so that this model takes center stage in this literature review.

Although the need for a valid and objective DC instrument is increasingly urgent in various educational contexts, this urgency is much more pronounced in *the Islamic School* environment. Islamic schools, including madrasas, face unique challenges in equipping students with digital competencies that are in line with Islamic moral, moral, and ethical values. In the midst of an unfiltered flow of information, issues such as digital moral literacy, security of interacting in cyberspace, and social media ethics are important issues that must be accommodated in the Islamic Religious Education (PAI) curriculum and strengthening religious character. Therefore, *the DC* measurement instrument used in Islamic educational institutions needs to be not only psychometrically accurate, but also sensitive to the context of Islamic values. The application *of the Many-Facet Rasch Model (MFRM)* is becoming increasingly relevant because it can capture the complexity of assessments involving cognitive, affective, and religious-based ethical behavior dimensions, including the variability of appraisers that often occur in observation-based assessments in Islamic educational institutions. Thus, this

research not only offers a methodological contribution to the global literature, but also provides an empirical foothold to improve the quality of digital competency evaluation in Islamic schools that are adapting to the challenges of the digital era.

Therefore, the biggest gap in the literature is the absence of a global systematic synthesis that maps: (1) the extent to which MFRM has been adopted in *the analysis of DC instruments*, (2) what methodological trends have emerged from this use, and (3) the implications of bibliometric findings on the development of *DC* instruments in *the Islamic School* environment. A comprehensive bibliometric review is needed to unify these three research variables, provide a clear *knowledge map*, and identify future directions of exploration.

Based on the identified literature gaps, this study has a clear distinction and novelty. In contrast to traditional literature reviews that are only narrative in nature or isolated MFRM application studies, this study adopts a global bibliometric approach. The bibliometric methodology allows researchers to provide an objective quantitative and visual analysis of intellectual structures, collaboration trends, and dominant research themes related to the application of MFRM on *DC instruments* which will later become a reference to be applied to Islamic educational institutions. The main novelty lies in the provision of a comprehensive map of knowledge, identifying key players (researchers, institutions, journals), and mapping the evolution of this topic from its inception to the present day.

This study is expected to be able to make an important contribution in various aspects. From a methodological perspective, this study offers empirical guidelines that can be a reference for other researchers who want to apply MFRM in assessing digital competencies in Islamic educational institutions. Theoretically, this study seeks to bridge the two domains that have been developing separately, namely modern psychometrics and *DC* studies, through a systematic synthesis of literature. Meanwhile, on the practical aspect, this study presents various best practices in the process of validating *DC* instruments used in various countries, so that it can be a valuable consideration for educators and policy makers. To achieve this goal, this study seeks to answer *the following main* research questions:

1. What are the trends in publication and scientific growth related to the application of *the Many-Facet Rasch Model (MFRM)* to *the* analysis of DC instruments globally?
2. What are the main *research themes* and emerging *topics* that are dominant in the literature that photographs MFRM and *DC instruments*?
3. How can these bibliometric findings have specific implications for the analysis and development of *DC* instruments in the *Islamic School environment*?

## B. Research Methods

This section systematically describes the methodology used in this bibliometric review study, which aims to map the intellectual structure and global trends related to the application of *the Many-Facet Rasch Model (MFRM)* in *the analysis of DC* instruments. This study uses a descriptive bibliometric analysis design to provide a quantitative and objective evaluation of the patterns of publication, collaboration, and thematic evolution of the relevant corpus of literature. As Donohue states, bibliometric analysis is a quantitative method for analyzing data from scientific publications to measure the quality, impact, and trends of research, as well as to map the relationships between researchers, institutions, and fields of science, as well as to identify the most influential research (Donohue, 1972; Huang et al., 2019; Ma & Ismail, 2025) The following in figure 1 are presented the steps of bibliometric analysis (He & Li, 2025).

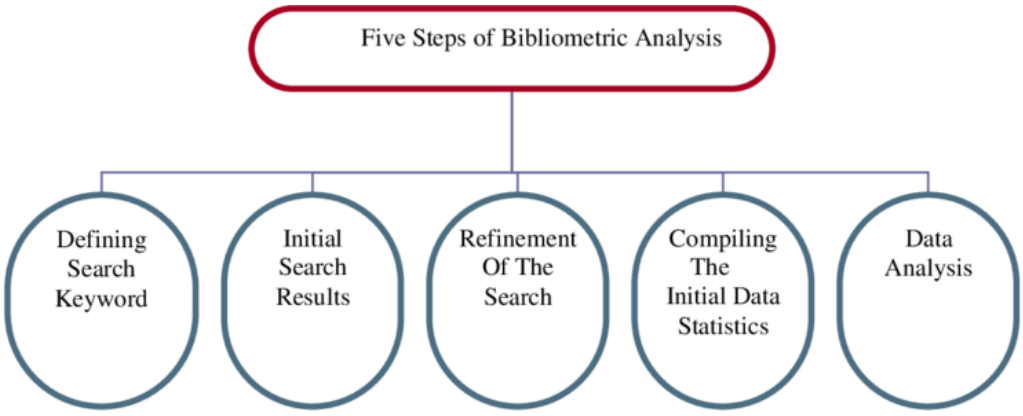


**Figure 1.** Bibliometric Analysis Steps

All data used in this analysis are exclusively sourced from the Scopus database, which is recognized for its comprehensive coverage and high-quality indexing of international publications. Scopus was chosen because it is the largest database of references and citations covering leading journals globally, which is ideal for bibliometric analysis. The literature search was conducted in November 2025 with no preliminary year limitation, allowing comprehensive coverage of Rasch-based research before applying relevance-based filtering at later stages. The search strategy was developed by combining *keywords* that reflect two main components, namely advanced psychometric methodology in this case MFRM and the context of the application in this case the measurement *of digital citizenship*. A structured search strategy using Boolean operators was applied to the TITLE-ABS-KEY fields.

The main query string used was: "Rasch model" AND ("character assessment" OR "moral education" OR "digital citizenship" OR "digital ethics" OR "online citizenship") AND ("Islamic school" OR "Islamic education").

The raw data extracted from Scopus then undergoes a rigorous screening process. The first stage involves restricting the type of document to be relevant to scientific research, namely Articles. Irrelevant document types, such as *reviews, conference papers*, *editorial materials, erratums,* or *notes*, are excluded. The next stage is manual screening, where the researcher reviews the title and abstract of each remaining entry to ensure the relevance of the topic, i.e. that the document must explicitly report on the *application* of MFRM or its variations to *measure* or *analyze* aspects related to Digital Citizenship. This process results in a final corpus of publications that serves as the basis for all quantitative analysis.

**Table 1.** Inclusion and Exclusion Criteria

| Aspect | Inclusion Criteria | Exclusion Criteria |
|---|---|---|
| **Document Types** | Article | *review, conference paper*, and other than articles |
| **Research Focus** | Multifaced Rasch Model | Rasch model in general |
| **Language** | English | In addition to English language articles |
| **Document Access** | Open access | Gold, green, bronze, hybrid gold |

Following the search strategy and eligibility criteria, the study selection process was conducted systematically using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework to ensure transparency and reproducibility in identifying relevant literature. The identification, screening, eligibility, and inclusion stages are illustrated in the PRISMA flow diagram (Figure 2), which presents the number of records retrieved, excluded, and retained at each step of the review process.

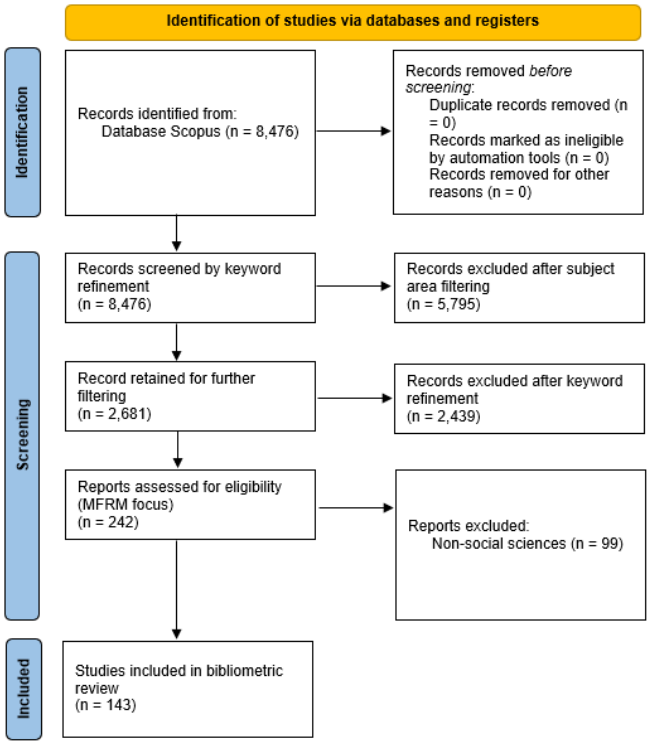


**Figure 2.** *Diagram Flow PRISMA Summary*

Figure 2 presents the PRISMA-based selection process. The Scopus search identified 8,476 records using the keyword "Rasch Model." After subject-area filtering, 2,681 records remained. Further keyword refinement focusing on MFRM reduced the dataset to 242 reports. After eligibility assessment, 143 studies met the inclusion criteria and were included in the final bibliometric analysis.

Data analysis was carried out entirely using *R-Studio* software by utilizing the bibliometrix library. *R-Studio* was chosen because of its strong capabilities in big data management and its comprehensive functionality for bibliometric analysis. The analysis is divided into several categories: (1) Performance Descriptive Analysis: Includes the calculation of annual publication trends, identifying the most productive journals, authors, and countries on this topic. (2) Citation Analysis: Determine the most influential documents based on the number of global citations and normalization. (3) Collaborative Network Analysis: Mapping author and country networks through *co-authorship* analysis to understand the scientific social structure of this field. (4) Thematic Analysis: Using *the co-occurrence* analysis of *Author Keywords* and *Keywords Plus* to identify dominant theme clusters and visualize the conceptual *structure of* the literature. The results of this analysis will be the basis for answering all the research questions that have been set.

## C. Results and Discussion

### Results

This section presents the results of a bibliometric analysis conducted on the global literature on the application of the Many-Facet Rasch Model (MFRM) in the context of Digital Citizenship instrument analysis. Based on the search strategy that has been established in the Methods Chapter, the initial data will be extracted from the Scopus database on November 19, 2025. After being extracted in *Comma Separated Values* (CSV) format, the data is imported into the R-Studio environment and processed using a bibliometrix package. After a data cleansing process that included the removal of duplication and filtering by document type and topic relevance, a final corpus of 143 publications was produced. It is this corpus that forms the quantitative basis for all bibliometric analyses presented in the following sub-chapters. The results of this study will be divided based on the type of analysis conducted, starting from publication performance trends, to mapping collaboration networks and thematic structures.

### Analysis of the Growth Trend of Scientific Publications Related to MFRM

The growth trend of scientific publications related to the application of the Many-Facet Rasch Model (MFRM) shows a fluctuating but significant increase throughout the study period from 1997 to 2025. Figure 2 visualizes the production of annual articles sourced from the Scopus database.

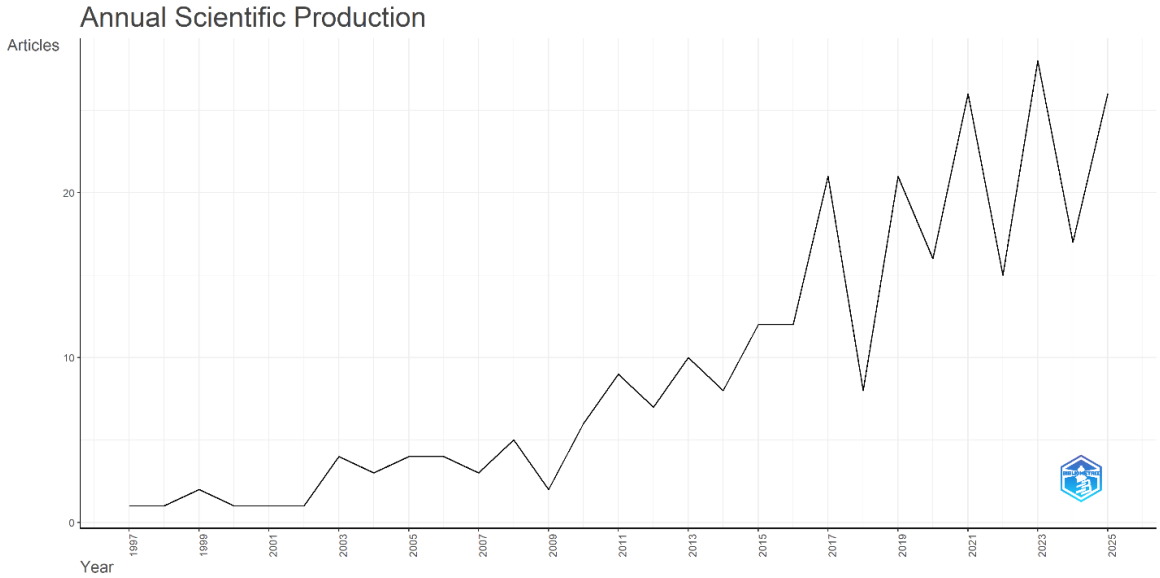


**Figure 3.** Annual Article Production
Source: Bibliometric Analysis Results from RStudio Application

In general, the literature in this field can be divided into two main phases:

1. Early Phase and Stagnation (1997-2009): During this period, scientific production was very limited. After the emergence of the first publications in 1997 and a small increase in the late 1990s (1999-2000), there was a significant period of stagnation in the early 2000s (2000-2002) in which there were few or no published articles. Annual production moves in a very low range, generally under 5 articles per year. In 2008 there were 5 articles at most. This indicates that the use of MFRM in this context is still a novel idea or limited to a very specific research community.
2. Acceleration and High Fluctuation Phases (2010-2025): After 2009, the field experienced substantial acceleration. Annual production shows a clear upward trend, breaking the 10-article mark per year consistently since 2015. This improvement reflects a growing recognition of the need for advanced psychometric models, such as MFRM, to validate increasingly complex digital instruments. Despite the increasing trend, this period was marked by sharp fluctuations, demonstrating the sensitivity of this field to research and funding cycles. The highest peak occurred in 2023 with production reaching 28 articles, followed by a sharp decline in 2024, but again showing a strong increase in the partial data for 2025. These sharp fluctuations, especially after 2017 in which publications reached 21 articles, underscore the nature of research that is project-based and not evenly distributed from year to year.

**Conceptual Structure Analysis: Keyword *Co-occurrence* Networks**

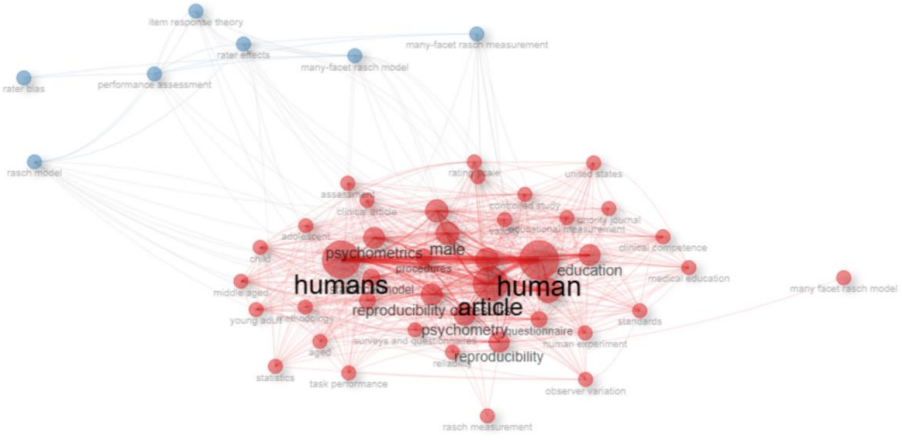


**Figure 4.** Keyword *Co-occurrence* Network
Source: Bibliometric Analysis Results from RStudio Application

Figure 4 shows the conceptual structure of the literature analyzed through co-occurrence mapping of author keywords. The keywords that often appear simultaneously describe the proximity of the research theme, forming a number of colored clusters that represent the main focuses in this field. In general, it is seen that two large clusters dominate the network. The first cluster, colored blue and located in the upper left, represents an advanced methodological theme. It contains keywords such as *Many-Facet Rasch Model*, *Item Response Theory*, *rater bias*, *rater effects*, and *performance assessment*.

This set of keywords shows the researchers' great concern for the issue of validity and reliability in performance appraisals, especially as it relates to differences between raters. The strong linkages between MFRM, IRT, and rater bias suggest that this approach was chosen to address subjective variation in assessment, something that is very much in line with the context of performance-based Digital Citizenship evaluation. Meanwhile, the second cluster, which appears dominant in the middle to the bottom of the network, which is marked in red, shows a broader theme of application as well as attention to psychometric quality in general. This cluster includes keywords such as *humans*, *psychometrics*, *reproducibility*, *reliability*, and *education*. This pattern shows that MFRM is most widely used in educational studies and research involving human participants, with an emphasis on ensuring reliable and replicable instruments. The presence of *Rasch Model* keywords in general in this cluster also indicates that MFRM is seen as an integral part of the broader framework of Rasch theory.

**Country Collaboration *Network***

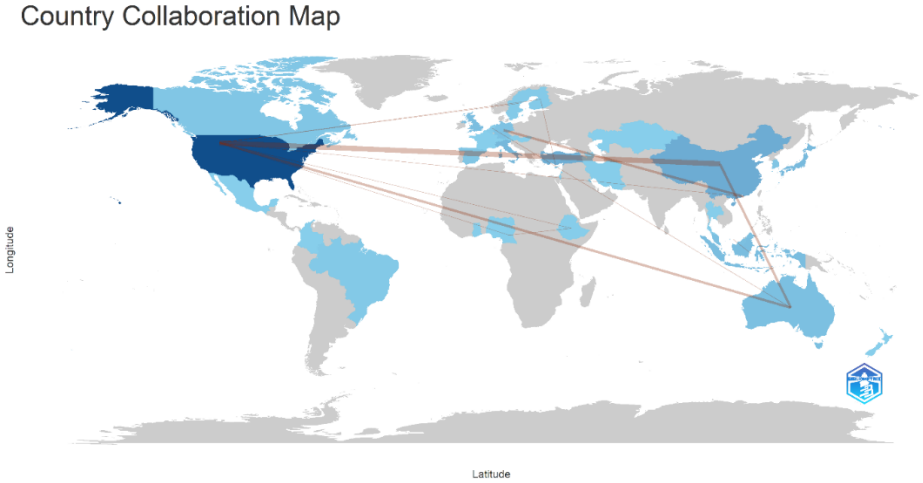


**Figure 5.** State Collaboration Network
Source: Bibliometric Analysis Results from RStudio Application

Figure 5 illustrates the pattern of international collaboration in research related to the Many-Facet Rasch Model (MFRM) and the analysis of digital instruments through a world map showing productivity levels and employment relations between countries. Blue of different intensities marks the level of contribution of the publication, while the connecting lines indicate the direction and strength of collaboration. From this visualization, it is clear that the United States is the center of academic activity, characterized by the darkest blue color and the main point of various global collaboration paths. Other countries such as the United Kingdom, Germany, the Netherlands, China, South Korea, Australia, and Canada also show a fairly high level of engagement, reflecting the concentration of MFRM research in regions that have a strong tradition in psychometrics and educational technology.

The collaborative network shows that the United States plays a global hub with intense relations with China, Australia, and a number of European countries. The thickness of the lines on the pathways signifies a high frequency of co-authorship, which in turn supports the exchange of advanced analytical methodologies, including the application of MFRM in a variety of cross-cultural digital contexts. The strong pattern of collaboration between the U.S. and countries in the Asia-Pacific region underscores the shared interest in improving the reliability and objectivity of measurement in a diverse digital environment.

From Indonesia's perspective, the map shows that Indonesia has been included in the collaboration network, although the intensity of the blue color is still relatively moderate compared to the central countries. This shows that Indonesia has a contribution but is not yet a major player in this field. However, the existence of a line of collaboration connecting Indonesia with countries such as the United States and Australia indicates that Indonesian research has begun to be integrated into the global scientific conversation. This situation implies great potential for expanding international cooperation, especially in the development of digital instruments and the use of MFRMs as a more objective and sensitive measurement approach to cultural variations. With the increasing need for digital competency measurement in Indonesia, more intense engagement in this global research network can be a strategic opportunity to strengthen academic quality and the relevance of national education policies.

**Contribution of Correspondence Authors and Collaboration Patterns**

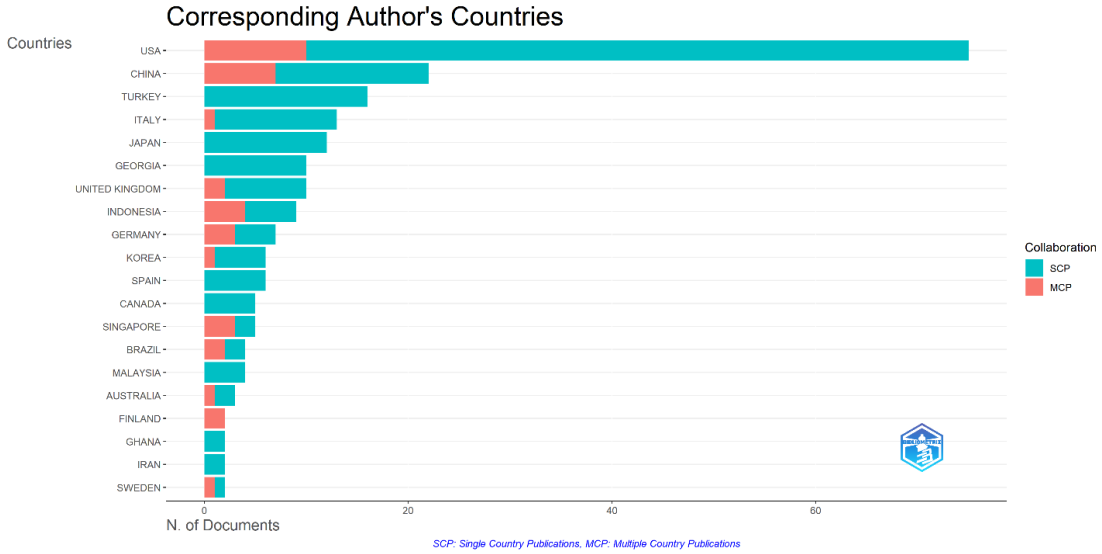


**Figure 6.** Contribution of Correspondence Authors and Collaboration Patterns
Source: Bibliometric Analysis Results from RStudio Application

Figure 6 shows the ranking of countries based on the total number of publications counted through the affiliation of the correspondence authors, while also showing their collaboration patterns through the proportion of single publications (SCPs) and collaborative publications (MCPs). From these visualizations, it is clear that the United States dominates scientific production in research related *to the Many-Facet Rasch Model (MFRM)* and digital instrument analysis. With more than 70 publications, the U.S. contribution far surpasses that of other countries such as China and Turkey, which recorded a total of 22 and 18 publications, respectively. This gap indicates that the US is becoming a major center of innovation as well as a global reference in the MFRM methodology. The SCP and MCP patterns also show different research approaches between countries; U.S. publications do come largely from independent efforts, but the proportion of international collaborations remains large and in line with its position as a global collaboration hub. Meanwhile, China is still dominated by SCP publications despite starting to build international collaboration, and Italy is seen to be more active in cooperating with foreign researchers. In contrast, Turkey and Japan produce more independent publications.

In the Indonesian context, the graph shows that Indonesia occupies the eighth position with about ten publications, an achievement that places it above developed countries such as Germany, Korea, and Canada. Interestingly, most of Indonesia's publications are the result of international collaboration, as can be seen from the

dominance of red bars representing MCP. This pattern suggests that Indonesian researchers are making greater use of global networks to access methodological expertise, especially in advanced psychometrics such as MFRM. This approach is very strategic because it provides opportunities to improve the quality of research as well as the visibility of publications in reputable databases such as Scopus. Overall, the findings in Figure 5 not only confirm the position of the United States as an MFRM research center, but also highlight that Indonesia, while not yet a major player in volume, has demonstrated a significant role through strong international collaboration.

**Intellectual Structure: Analysis of *Three Field Plots***

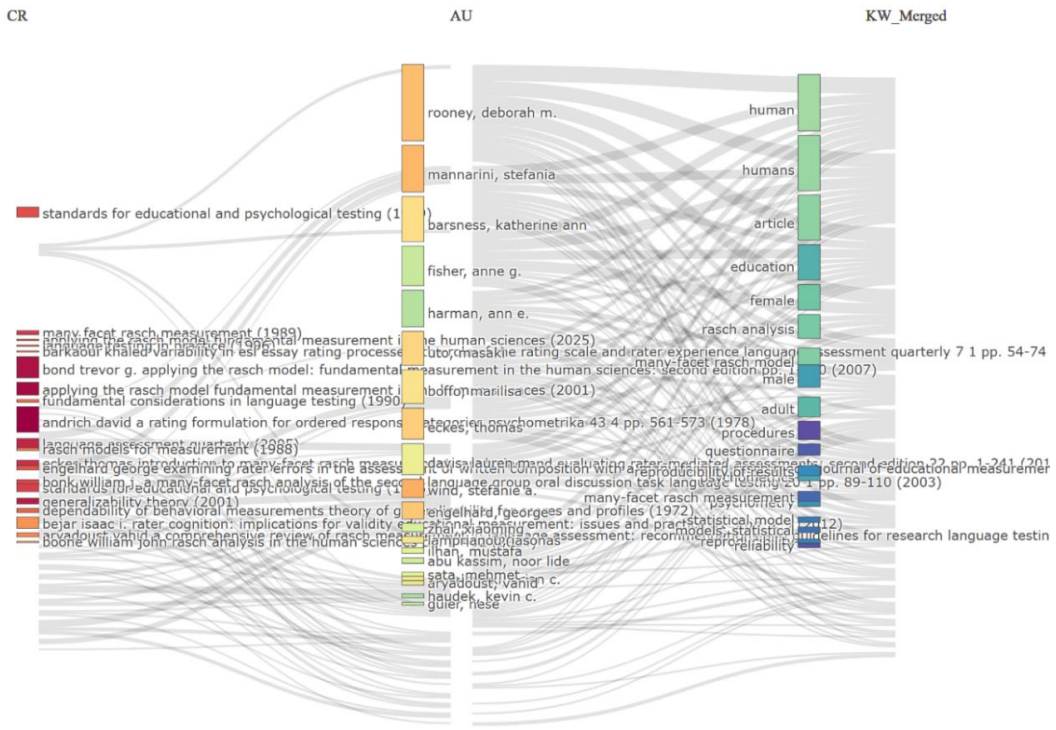


**Figure 7.** Three *Field Plot Analysis*
Source: Bibliometric Analysis Results from RStudio Application

Figure 7 shows the intellectual structure of this field of research by connecting three main components: theoretical foundations (*Cited References (CR),* key authors (*Authors/AU),* and thematic focus (*Merged Keywords/KW_Merged*). The width of the connection line indicates the strength of the relationship between the elements. In the field of theoretical foundations, the literature related to MFRMs and digital

instruments relies heavily on fundamental works in psychometrics and Rasch's theory, including references to methodological standards such as *Standards for Educational and Psychological Testing* as well as research on Generalizability Theory and dependability of behavioral measurements. In addition, there is a strong foothold in the works that build on the basis of the Rasch Model and the Many-Facet Rasch Model, showing a close connection with Item Response Theory (IRT). In the field of key authors, individuals such as Rooney, Mannarini, Barsness, Fisher, and Harman emerged as key actors connecting theory with thematic applications, acting as bridges that transferred methodological knowledge to the focus of research. Meanwhile, thematic focus areas highlight dominant keywords such as Human/Humans, Article, Education, Female, Rasch Analysis, and Many-Facet Rasch Model, which affirm the direction of research in the context of human behavior and education.

Connection line analysis shows a clear flow of knowledge: from theory to author, and from author to theme, showing how methodological foundations are consistently applied by prolific authors into human-oriented research and education. The thick relationship that associates the Rasch/MFRM reference with the keyword methodology emphasizes that this study does not only use the method passively, but also discusses and develops its application in depth. Overall, these Three Field Plots represent a mature research ecosystem, guided by the basic literature, driven by prolific authors, and focused on the application of advanced methodologies in the context of the humanities and education.

**Thematic Evolution Analysis: Topic Trends**

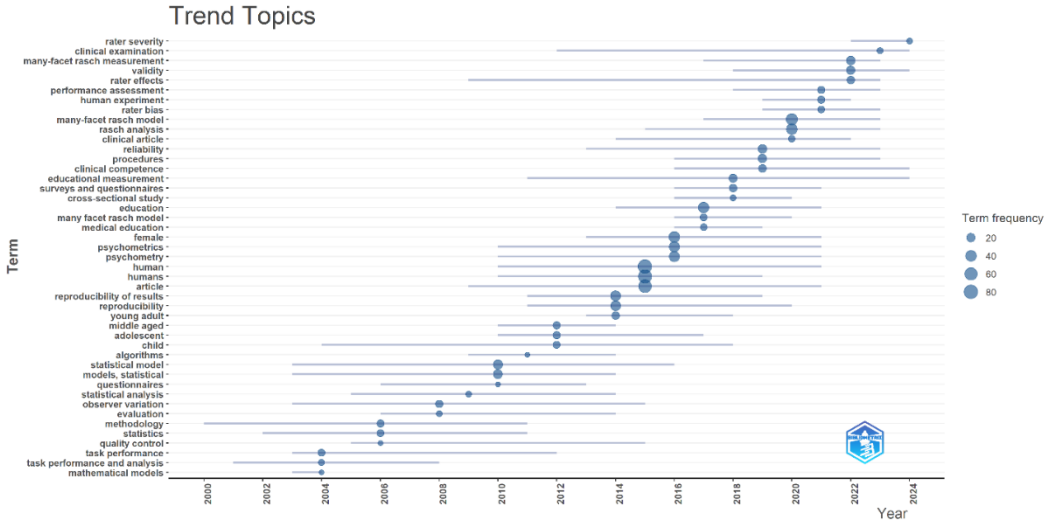


**Figure 8.** Trending Topics
Source: Bibliometric Analysis Results from RStudio Application

Figure 8 illustrates the temporal trends of the most relevant keywords in the literature, showing the periods of occurrence of a particular term as well as its intensity through the size of the bubble. This analysis helps identify well-established topics that arise in the field of research. In the early phases (before 2010), the dominant topics were methodological and fundamental, including terms such as *Evaluation*, *Methodology*, *Statistics*, *Statistical Analysis*, *Models*, and *Observation*, which showed a focus on the development of basic statistical frameworks before MFRM was widely implemented. In addition, the issue of quality control also emerged early, with keywords such as *Quality Control* and *Task Performance*, indicating an early awareness of the reliability of performance measurement.

The period 2010–2018 marked a surge in interest in the context of application and instrument quality issues. Reproducibility-related terms, such as *Reproducibility* and *Reproducibility of Results*, reached their peak in popularity between 2014–2018, confirming the researcher's focus on the validity and replicability of measurement results. In addition, the keywords related to the research subjects *Child*, *Adolescent*, *Young Adult*, *Middle Aged*, *Female*, and *Male* experienced a significant increase around 2012–2016, indicating the expansion of testing instruments in various demographic groups to reduce potential bias. The high activity on the keywords *Education* and *Educational Measurement* in 2016–2018 also confirms that the field of education remains the main application context of this literature.

The most recent phase, post-2018, shows a shift back to sophisticated methodological issues. Terms such as *Rater Severity*, *Rater Bias*, *Rater Effects*, and *Performance Assessment* show a significant increase in frequency, signaling the research community's attention to specific facets and evaluator bias issues, which can be addressed with MFRM. The terms *Many-Facet Rasch Measurement* and *Validity* also indicate high peaks of activity, confirming the role of MFRM as the primary methodological approach in addressing rater bias issues. In addition, the emergence of keywords such as *Clinical Examination* and *Clinical Competence* signifies the expansion of MFRM applications into the clinical realm, which demands accuracy and objectivity on par with educational and digital contexts. Overall, this temporal trend illustrates the evolution of research from a basic methodological focus in the early 2000s, to applications and demographic issues in the mid-2010s, to an emphasis on specific methodological aspects related to facets and rater bias in recent periods, reinforcing the importance of MFRM as a current solution.

**Thematic Brief Visualization: *Word Cloud***


**Figure 9.** Word Cloud
Source: Bibliometric Analysis Results from RStudio Application

Figure 9 shows a word cloud that summarizes the main thematic focus of the entire corpus of literature, where word size reflects the frequency of its occurrence and provides a visual representation of the centrality of the topic. This analysis confirms that this research is dominated by three main axes. First, in terms of methodology, the *phrase many-facet rasch model* stands out the most, confirming that MFRM is the main methodological focus. Along with that, related terms such as *rasch analysis*, *rasch model*, *psychometrics*, and *psychometry* also appear frequently, indicating strong research roots on Item Response Theory (IRT) and attention to high-quality measurement standards. Second, on the subject side, the keywords *human* and *humans* occupy a significant position after MFRM, emphasizing that the main application of this model is aimed at measuring human attributes. Third, regarding quality issues, terms such as *reproducibility*, *reliability*, and *validity* stand out, indicating that the main purpose of using MFRM is to improve and verify the psychometric quality of instruments, especially in the context of performance assessment and surveys.

In addition, word cloud analysis also highlights a secondary contextual focus. The terms *education* and *educational measurement* affirm the dominance of the field of education as a context of application, while the emergence of *clinical competence* and *medical education* indicates the expansion of the domain into the clinical domain. From a demographic perspective, keywords such as *female*, *male*, *adult*, and *adolescent* indicate attention to bias issues based on subject groups. The presence of the words *article* and *priority journal* also indicates that the literature analyzed

comes from highly reputable journals. Overall, this word cloud provides clear visual evidence that the research in this corpus focuses intensively on the application of MFRM to measure human attributes, with the primary goal of achieving high reliability and validity in the context of education and performance evaluation.

**Analysis of Publication Growth Accumulation by Country**

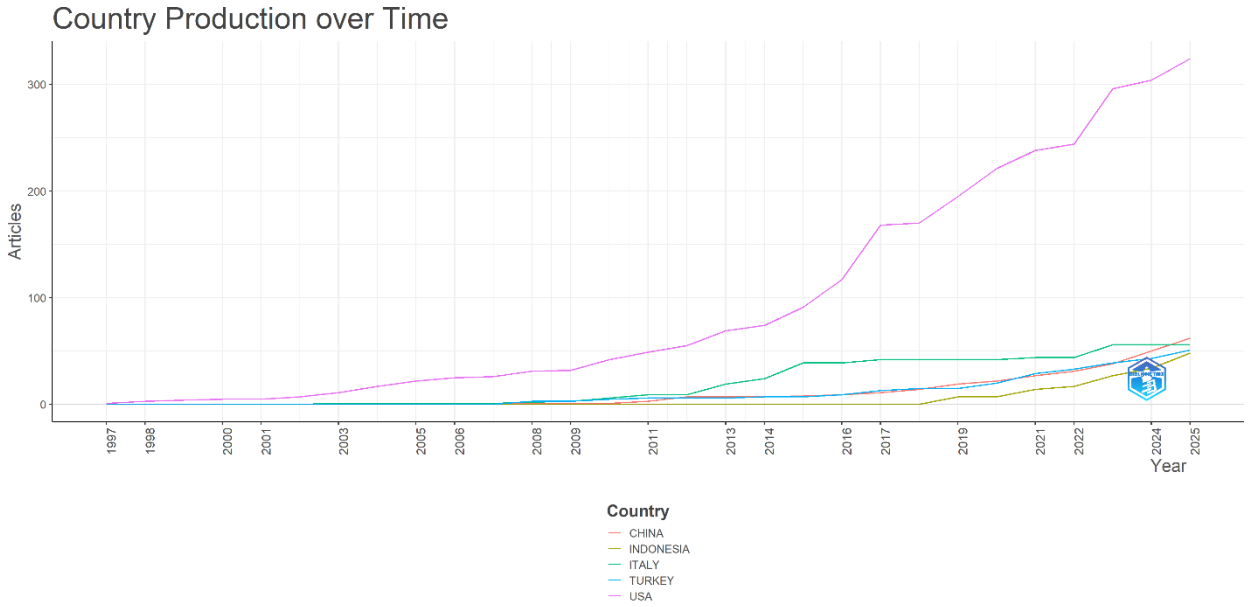


**Figure 10.** Analysis of Publication Growth Accumulation by Country
Source: Bibliometric Analysis Results from RStudio Application

Figure 10 presents a visualization of the cumulative publication production of the top five most relevant countries in the Many-Facet Rasch Model (MFRM) study in the digital context, namely the United States (USA), China (China), Turkey (Turkey), Italy (Italy), and Indonesia, throughout the study period. From the graph, it is clear that the dominance of the United States, depicted with a purple line, as the absolute leader in the number of publications. The production of US publications started early and showed steady and sharp growth, until by the end of the study period (2024–2025) it reached more than 300 articles, far exceeding the combined total publications of other countries. This confirms the role of the US as the initiator as well as the main contributor of knowledge in the field of MFRM applied to the digital context.

Meanwhile, other countries that are key players show slower cumulative growth rates. China (red) and Turkey (blue) showed significant acceleration especially after 2010, with a cumulative total of 60 articles at the end of the period,

although still well below the US. Italy (green) showed a rapid surge in publications in the mid-2010s (around 2013–2017), before growth flattened and picked up a bit again, signaling a period of intensive research.

Indonesia (yellow) occupies a relatively new position as a contributor. The growth of his publications began to be seen significantly in the last years of the study period, which was around 2018 and above. Although Indonesia's cumulative number is still the lowest among the five countries, at around 15 articles, its growth line shows a fairly steep acceleration in 2022–2025, reflecting the rapid increase in research interest and activity. Overall, this cumulative analysis confirms the central role of the United States as a global leader while highlighting Indonesia as an emerging player that is just beginning to actively expand its contribution to MFRM literature and the development of digital instruments.

## Discussion

### Trends in Publication and Scientific Growth

The results of the bibliometric analysis conducted in this study show significant scientific growth dynamics related to the application of the Many-Facet Rasch Model (MFRM) in the development and validation of *DC* instruments, especially in the context of education. However, despite this notable progress, many existing assessment practices still rely on traditional measurement approaches such as Classical Test Theory (CTT), which tend to produce less accurate results and are highly dependent on test-taker characteristics, rater subjectivity, and item difficulty variations (Maimunah et al., 2025)**.** This methodological limitation highlights the need for more robust measurement models such as MFRM, which can account for multiple facets simultaneously and provide fairer, more objective, and psychometrically defensible evaluations. The sharp increase in publication growth after 2010 shows that there is a growing concern for the need for advanced psychometric models that are able to provide objective evaluation of performance-based assessments in the digital environment (Chen et al., 2021). Assessment plays a crucial role in education, serving as a means to evaluate learners' performance, inform instructional planning, and assess the effectiveness of teaching (Mannan et al., 2025).

This increasing trend of publications indicates that the use of instruments to measure digital competence requires a methodological approach capable of addressing assessor bias and category variability, as affirmed by (Myford & Wolfe, 2004) that MFRM allows researchers to identify and correct differences in assessor rigidity as well as score inconsistencies. In the context of digital citizenship assessments, the ability to manage assessor bias is very important because

assessments are not solely based on written tests, but rather students' real performance in complex digital situations, including ethical interactions, online safety, and social responsibility (Choi et al., 2017). If the assessment process is effective, then educational institutions are considered successful in instilling moral values and forming students with personalities and noble character through learning activities (Alafthoni, 2024).

Bibliometric findings that show the dominance of publications from the United States confirm that methodological innovations in the field of modern psychometrics are growing rapidly in countries with strong research traditions in the measurement of education (Goodwin, 2016). The position of the United States as a center of global collaboration shows that the development of MFRM-based assessment instruments is not only methodological but also strategic in supporting the quality of digital education policies (Morales-Álvarez et al., 2025). On the other hand, Indonesia's involvement in showing an increasing trend of publications through international cooperation strengthens the statement that academic collaboration can be a means of increasing national research capacity and transforming the quality of educational assessment (Effendi et al., 2020). One of the interesting developments in the transformation of the quality of educational assessment that has been widely discussed in recent research is the application of authentic assessment and technology-based diagnostics, which are able to create a more personalized and interactive learning experience (Fadlillah & Kusaeri, 2024). This condition is a valuable opportunity for Islamic madrassas and schools to strengthen a more valid and scientific evidence-based digital evaluation approach through a global scientific network.

**Main *Research Themes* and Emerging *Topics***

The analysis of the thematic structure reveals two major focuses of the study, namely the methodological fields related to rater bias, psychometrics, and reliability, as well as the applicative fields related to education, human performance, and instrument validity. This thematic structure is consistent with the view (Pangrazio & Sefton-Green, 2021) that today's digital education is no longer limited to technology literacy, but involves ethics, safety, and moral responsibility. In this context, the existence of the MFRM model is very relevant because the digital citizenship measurement instrument must be able to capture the actual behavior of students in digital situations that are risky and require ethical decision-making (Reilly et al., 2024). The use of conventional psychometric models is often unable to address bias between raters, so the application of MFRM provides a more comprehensive solution because it considers the variance of items, respondents,

raters, and rating categories simultaneously (John & Linacre, 1993). The problem of assessor bias is one of the biggest issues in digital performance assessment, as revealed (Arifiyanti et al., 2023) that gender bias and academic bias can arise when assessors have different levels of rigidity in assigning scores.

**Implications for the Analysis and Development of *DC* Instruments in The *Islamic School***

In the context of Islamic schools, the interpretation of bibliometric findings has important implications. Madrasas and schools with Islamic identity have a moral responsibility to ensure that digital character assessments and media ethics reflect Islamic values such as justice, trust, and responsibility (Iskandar et al., 2025). For this reason, the objectivity of assessment is a central principle, especially when the assessment is carried out using performance rubrics that are vulnerable to the assessor's subjectivity (Dogan & Uluman, 2017). The implementation of MFRM helps ensure that assessments of students' digital morals truly reflect their actual abilities, not biased perceptions of assessors on demographic identity, school background, or intensity of religiosity. This is in line with the view (Nordin et al., 2016) that the quality of *DC instruments* should be tested not only technically, but also through a cross-cultural and value-based validity approach.

In addition, findings regarding global scientific collaboration networks show that the growth of Indonesia's contribution based on international cooperation can open up space for innovation in the development of digital citizenship instruments based on Islamic values. Faith-based madrassas and schools need to develop authentic assessment instruments that can measure students' ability to apply moral values in a digital environment, including digital trust, information sharing ethics, virtual empathy, and professionalism in online communication (Connolly & Miller, 2022). Through the implementation of MFRM, researchers and educators can ensure that digital moral performance assessments are conducted fairly and without bias, so that the results of the assessment can be used to design more effective and systematic learning interventions (Mclaughlin et al., 2023).

The implications of the bibliometric findings also show the urgent need to develop *DC* instruments that integrate Islamic values of religiosity and morality, especially in an era when learners face digital pressures such as cyberbullying, disinformation, abuse of privacy, and a culture of social media hedonism (Naeem & Mushibwe, 2025). In the Islamic view, the existence of assessment instruments that can measure digital morality is not only a pedagogical goal but also worship and the protection of human dignity. Therefore, this study emphasizes the need for an MFRM-based evaluation approach to help Islamic schools assess digital competence

and religious character in a balanced manner. In alignment with this need, revising the Islamic Religious Education curriculum becomes essential through the integration of learning modules that promote a comprehensive understanding of universal values such as peace, tolerance, and respect for cultural diversity. Such a curriculum design will not only strengthen students' digital ethics but also cultivate socially responsible behavior rooted in Islamic moral teachings (Fitriani & Hilmy, 2023).

Thus, the results of this bibliometric research show that the application of MFRM provides a strong methodological and practical basis to improve the quality of *DC evaluation* in Islamic schools through objective, fair, and value-oriented performance assessments. These findings explicitly answer the research question that MFRM has become a key approach in the global literature to address assessment bias, build the validity of digital instruments, and expand its application to value-based education contexts. This scientific narrative serves as an argumentative basis to recommend further research on the integration of Islamic values in performance-based digital assessment with the support of the MFRM methodology.

## D. Conclusion

This study aims to comprehensively map the global development of the use of the Many-Facet Rasch Model (MFRM) in the analysis of Digital Citizenship instruments and identify its implications for the context of Islamic education. The bibliometric results show that MFRM-related publications have increased significantly in the last two decades, reflecting the strong need for a psychometric approach that is able to ensure objectivity and fairness in digital performance-based assessments. The main contribution of this research lies in the provision of a knowledge map of key actors, global scientific collaboration, and the thematic direction of developing research, while emphasizing the relevance of the application of MFRM to strengthen the quality of digital competency evaluation in Islamic schools that demand assessments based on moral values, trust, and justice.

This study suggests the importance of developing Digital Citizenship assessment instruments that are integrated with Islamic values and tested using a multifaceted approach to minimize assessor bias. Further studies are recommended to explore the empirical application of MFRM to digital moral assessment in madrasas and conduct cross-cultural Differential Facet Functioning (DFF) analysis as well as expand international collaboration in evidence-based assessment innovations.

**References**

Alafthoni, F. Z. (2024). Evaluation of Islamic Religious Education and Character Learning Using the Countenance Stake Model. *Edukasi Islami: Jurnal Pendidikan Islam*, *13*(01), 59–72. https://doi.org/10.30868/ei.v13i01.6115

Arifiyanti, F., Soeharto, S., Amukune, S., Nguyen, S. Van, Aburezeq, K., Hidayatullah, A., & Sarimanah, E. (2023). Heliyon Investigating rater – student interaction , gender bias , and major bias in the assessment of research seminar presentation. *Heliyon*, *9*(6), e16548. https://doi.org/10.1016/j.heliyon.2023.e16548

Aslanoglu, A. E., Karakaya, I., & Sata, M. (2020). Evaluation of University Students' Rating Behaviors in Self and Peer Rating Process via Many Facet Rasch Model. *Eurasian Journal of Educational Research*, *89*, 25–46. https://doi.org/10.14689/ejer.2020.89.2

Bahrouni, F. (2016). Using Multi-Facet Rasch Model (MFRM) in Rater-Mediated Assessment. *The Journal of Teaching English for Specific and Academic Purpose*, *4*(53 5), 195–212.

Brennan, R. L. (2011). Applied Measurement in Education Generalizability Theory and Classical Test Theory. *Applied Measurement in Education*, *August 2013*, 37–41. https://doi.org/10.1080/08957347.2011.532417

Chen, L. L., Mirpuri, S., Rao, N., & Law, N. (2021). Conceptualization and measurement of digital citizenship across disciplines. *Educational Research Review*, *33*(January), 100379. https://doi.org/10.1016/j.edurev.2021.100379

Choi, M., Glassman, M., & Cristol, D. (2017). Computers & Education What it means to be a citizen in the internet age : Development of a reliable and valid digital citizenship scale. *Computers & Education*, *107*, 100–112. https://doi.org/10.1016/j.compedu.2017.01.002

Clayson, P. E., Carbine, K. A., Baldwin, S. A., Olsen, J. A., & Larson, M. J. (2021). Using generalizability theory and the ERP Reliability Analysis ( ERA ) Toolbox for assessing test-retest reliability of ERP scores part 1 : Algorithms , framework , and implementation. *International Journal of Psychophysiology*, *January*. https://doi.org/10.1016/j.ijpsycho.2021.01.006

Connolly, R., & Miller, J. (2022). Evaluating and Revising the Digital Citizenship Scale. *Informatics*, 1–13.

Dass, M. A., & Kumar, M. P. M. P. (2025). Critical Digital Citizenship: a Scale Development and Validation Study. *International Journal of Technology and Design Education*, 1493–1513.

Dogan, C. D., & Uluman, M. (2017). A Comparison of Rubrics and Graded Category Rating Scales with Various Methods Regarding Raters ' Reliability. *Educational Sciences: Theory and Practice*, 631–651. https://doi.org/10.12738/estp.2017.2.0321

Donohue, C. (1972). A Bibliometric Analysis of Certain Information Science Literature. *Journal of the American Society for Information Science*.

Effendi, M., Bin, E., & Matore, M. (2020). Rating Performance among Raters of Different Experience Through Multi-Facet Rasch Measurement ( MFRM ) Model. *Journal of Measurement and Evaluation in Education and Psychology*, *11*(2), 147–162. https://doi.org/10.21031/epod.662964

Fadlillah, N., & Kusaeri. (2024). Optimizing Assessment for Learning in Islamic Education through Authentic and Diagnostic Assessment : A Systematic Literature Review. *Jurnal Kependidikan: Jurnal Hasil Penelitian Dan Kajian Kepustakaan Di Bidang Pendidikan, Pengajaran Dan Pembelajaran*, *10*(2), 654–667.

Feriandi, Y. A., Budimansyah, D., & Komalasari, K. (2022). Developing measurement instrument of students ' citizenship ecological behavior on citizenship education course in middle school. *Journal of Human Behavior in the Social Environment*, *00*(00), 1–13. https://doi.org/10.1080/10911359.2022.2120590

Fitriani, E., & Hilmy, M. (2023). Evaluation of Islamic Religious Education : Optimal Approach to Increase Tolerance. *Edukasia : Jurnal Penelitian Pendidikan Islam*, *18*(2), 293–312. https://doi.org/10.21043/edukasia.v18i2.26618

Goodwin, S. (2016). Assessing Writing A Many-Facet Rasch analysis comparing essay rater behavior on an academic English reading / writing test used for two purposes. *Assessing Writing*. https://doi.org/10.1016/j.asw.2016.07.004

Guler, N. (2014). Analysis of Open-Ended Statistics Questions with Many Facet Rasch Model. *Eurasian Journal of Educational Research*, *55*, 73–90.

He, L., & Li, L. (2025). Future trends and challenges for citizenship ' s digital literacy

research : a bibliometric analysis. *Discover Education*.

Heru, W., Indah, B., Sa, H., Banu, N., Naidu, M., & Prasetyo, E. (2022). Using Rasch model analysis for assessing psychometric properties of digital citizenship in Indonesian students. *Journal of E-Learning and Knowledge Society*, *18*(3), 96–110.

Ho, P. (2019). A new approach to measuring Overall Liking with the Many-Facet Rasch Model. *Food Quality and Preference*, *74*(January), 100–111. https://doi.org/10.1016/j.foodqual.2019.01.015

Huang, C., Yang, C., Wang, S., Wu, W., Su, J., & Liang, C. (2019). Evolution of topics in education research : a systematic review using bibliometric analysis. *Educational Review*, *00*(00), 1–17. https://doi.org/10.1080/00131911.2019.1566212

Iskandar, R., Maksum, A., & Marini, A. (2025). Social Sciences & Humanities Open Digital citizenship literacy in Indonesia : The role of privacy awareness and social campaigns. *Social Sciences & Humanities Open*, *12*(June), 101697. https://doi.org/10.1016/j.ssaho.2025.101697

John, M., & Linacre, J. M. (1993). Generalizability Theory and Many-facet Rasch Measurement. *Annual Meeting of the American Educational Research Association*.

Jones, L. M., & Mitchell, K. J. (2015). Defining and Measuring Youth Digital Citizenship. *New Media & Society*. https://doi.org/10.1177/1461444815577797

Laura-de-la-cruz, K. M., Pérez-postigo, G., & Diaz-zavala, R. (2024). Digital Skills and Digital Citizenship Education: An Analysis Based on Structural Equation Modeling. *Journal of Technology and Science Education*, *14*(3), 738–755.

Ma, H., & Ismail, L. (2025). Bibliometric analysis and systematic review of digital competence in education. *Humanities and Social Sciences Communications*, *2025*, 1–17. https://doi.org/10.1057/s41599-025-04401-1

Maimunah, S., Hidayati, Z., & Mulyana, S. I. (2025). Comparison of GRM and GPCM in the Development of Higher Education Practice Assessment Instruments. *Scaffolding: Jurnal Pendidikan Islam Dan Multikulturalisme*, *7*(2), 286–303. https://doi.org/10.37680/scaffolding.v7i2.7281

Mannan, A., Suparto, & Kusaeri. (2025). Practices and Challenges of the Validity of Exploratory Factor Analysis ( EFA ) -Based Assessment Instruments : A Systematic Literature Review. *EDUKASIA: Jurnal Pendidikan Dan Pembelajaran*, *6*(1), 611–626. https://doi.org/10.62775/edukasia.v6i1.1463

Mclaughlin, J. E., Angelo, T. A., White, P. J., Peeters, M., & Green, K. E. (2023). Validating criteria for identifying core concepts using many-facet rasch measurement. *Frontiers in Education*, *April*, 1–8. https://doi.org/10.3389/feduc.2023.1150781

Morales-álvarez, A., Valdés-cuervo, A. A., Parra-pérez, L. G., García-vázquez, F. I., & Durón-ramos, M. F. (2025). Adolescents' appraisal and disclosure mediate the links between parental practices and adolescents' digital citizenship behaviors. *Humanities and Social Sciences Communications*, *2025*, 1–9. https://doi.org/10.1057/s41599-025-04485-9

Myford, C. M., & Wolfe, E. W. (2004). Detecting and Measuring Rater Effects Using Many-Facet Rasch Measurement : Part II. *Journal of Applied Measurement*, *5*(2), 189–227.

Naeem, K., & Mushibwe, C. P. (2025). Navigating digital worlds : a scoping review of skills and strategies for enhancing digital resilience among higher education students on social media platforms. *Discover Education*. https://doi.org/10.1007/s44217-025-00432-7

Nordin, M. S., Badariah, T., Ahmad, T., Zubairi, A. M., Ahmad, N., & Ismail, H. (2016). Psychometric Properties of a Digital Citizenship Questionnaire. *International Education Studies*, *9*(3), 71–80. https://doi.org/10.5539/ies.v9n3p71

Özüdog, G. (2025). Conceptualizing pre-service teachers' artificial intelligence readiness and examining its relationship with various variables: The role of artificial intelligence literacy, digital citizenship, artificial intelligence-enhanced innovation and perceived thr. *Information Development*, *41*(3). https://doi.org/10.1177/02666669251335657

Pangrazio, L., & Sefton-Green, J. (2021). Digital Rights , Digital Citizenship and Digital Literacy : What ' s the Difference ? *Journal of New Approaches in Educational Research (NAER Journal)*, *10*, 15–27. https://doi.org/10.7821/naer.2021.1.616

Reilly, M. O., Levine, D., Batchelor, R., Adams, S., Reilly, M. O., Levine, D., Batchelor, R., & Adams, S. (2024). Digital ethics of care and digital citizenship in UK primary

schools : Children as interviewers Digital ethics of care and digital citizenship in UK primary ABSTRACT. *Journal of Children and Media*, *18*(4), 585–604. https://doi.org/10.1080/17482798.2024.2394932

Sata, M., & Karakaya, I. (2020). Investigation of the Use of Electronic Portfolios in the Determination of Student Achievement in Higher Education Using the Many-Facet Rasch Measurement Model. *Educational Policy Analysis and Strategic Research*, *15*(1), 9. https://doi.org/10.29329/epasr.2020.236.1

Sebasti, J., & Lozano-d, A. (2021). Measuring Digital Citizenship : A Comparative Analysis. *Informatics*, 1–13.

Valentine, N., Durning, S. J., Shanahan, E. M., Der, C. Van, Schuwirth, L., Valentine, N., Durning, S. J., Shanahan, E. M., & Der, C. Van. (2022). The pursuit of fairness in assessment : Looking beyond the objective. *Medical Teacher*, *44*(4), 353–359. https://doi.org/10.1080/0142159X.2022.2031943

Veenhoven, R. (2012). Cross-national differences in happiness : Cultural measurement bias or effect of culture ? *International Journal of Wellbeing*, *2*, 333–353. https://doi.org/10.5502/ijw.v2.i4.4

Webster, J. (2025). Sociotechnical Imaginaries in a Postdigital World : Teachers ' Perceptions of Digital Citizenship Education. *Postdigital Science and Education*, *7*(3), 770–787. https://doi.org/10.1007/s42438-025-00557-w

Wolfe, E. W. (2015). Detecting and Measuring Rater Effects Using Many-Facet Rasch Detecting and Measuring Rater Effects Using Many-Facet Rasch Measurement : Part I. *Journal of Applied Measurement*, *February 2003*, 386–422.

Yuanyuan, P., Alias, B. S., Mansor, A. N., Rashid, M., & Hamid, A. (2025). Development and validation of the principals ' digital leadership instrument using Rasch measurement model. *International Journal of Evaluation and Research in Education (IJERE)*, *14*(3). https://doi.org/10.11591/ijere.v14i3.32214

Zlatkin-troitschanskaia, O., Schlax, J., Jitomirski, J., & Happ, R. (2019). Ethics and Fairness in Assessing Learning Outcomes in Higher Education. *Higher Education Policy*. https://doi.org/10.1057/s41307-019-00149-x