Jurnal Penelitian dan Evaluasi Pendidikan
Volume 28, No. 2, December 2024 (173-185)
Online: http://journal.uny.ac.id/index.php/jpep

Development of digital literacy assessment instrument for prospective teacher students in higher education

Hari Kantona*; Sudji Munadi
[1]Universitas Negeri Yogyakarta, Indonesia
*Corresponding Author. E-mail: hkantona@gmail.com

ARTICLE INFO

Article History
Submitted:
23 April 2024
Revised:
15 June 2024
Accepted:
30 September 2024

Keywords
digital literacy; instruments; prospective student educators

Scan Me:

ABSTRACT

This study aims to (1) develop an instrument to assess digital literacy skills among prospective teacher students, (2) examine its psychometric quality, and (3) describe students' digital literacy levels. The instrument development followed ten stages, including item writing, expert review, limited and large-scale testing, and data analysis. A total of 262 prospective teachers from UNY and UIN Sunan Kalijaga Yogyakarta participated in the large-scale test. Construct validity was tested using confirmatory factor analysis (CFA), and item characteristics were analyzed using the graded response model (GRM) within item response theory (IRT). Reliability was measured using Cronbach's Alpha. The final instrument contains 35 items across seven aspects: information literacy, digital scholarship, learning skills, ICT literacy, career and identity management, communication and collaboration, and media literacy. The instrument showed good psychometric quality based on Aiken's V, CFA, and IRT analysis. Measurement results showed 9% of students had excellent digital literacy, 19% good, 38% sufficient, 30% lacking, and 4% very lacking. These results highlight the need for targeted training and support to improve digital literacy among prospective teachers in Yogyakarta.

To cite this article (in APA style):
Kantona, H., & Munadi, S. (2024). Development of digital literacy assessment instrument for prospective teacher students in higher education. *Jurnal Penelitian dan Evaluasi Pendidikan, 28(2)*, 173-185 doi: https://doi.org/10.21831/pep.v28i2.72712

INTRODUCTION

The Indonesian government has introduced many new innovations to support the achievement of educational goals up to the present time. One of these is the Indonesian government, through the Ministry of Education and Culture in 2016, intensifying the National Literacy Movement (NLM) to cultivate a culture of literacy across all educational domains (Kemendikbud, 2017). The World Economic Forum in 2015 emphasized that mastery of six basic literacies, namely reading and writing literacy (language), numeracy, science literacy, digital literacy, financial literacy, and cultural citizenship literacy, is one of the 21st-century competencies required by all global citizens, especially learners (OECD, 2019; Wahyuningsih, 2021). Although the government intensified NLM in 2016, there are studies indicating that the digital literacy skills of Indonesian educators are still low (Nurlailah, 2022; Rahmawati, 2018; Setyaningsih et al., 2019). These studies have shown that educators' utilization of technology is limited to activities such as presentations and online assignments. This is due to educators' lack of knowledge on how to leverage existing technological developments to enhance the quality of learning. Therefore, as digital literacy skills are one of the literacies under NLM and based on these research findings indicating the low digital literacy skills of Indonesian educators, digital literacy still needs to be studied.

When observing educators' conditions during online teaching amid the pandemic, educators tend to be unprepared due to their lack of technological literacy. Therefore, educators

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v28i2.72712

require training to improve their digital literacy skills (Fitriani et al., 2023; Soeparno & Ismaniati, 2022; Warsiyah et al., 2022). This is because online teaching during the pandemic was carried out due to urgent circumstances, causing most educators to be unprepared for it (Sánchez-Cruzado et al., 2021). Additionally, based on research by Soeparno and Ismaniati (2022) explains the phenomenon of low digital literacy skills among educators starting from their college years. Furthermore, Falloon (2020), Nasrullah et al. (2017), and Rukayah et al. (2022) also state that digital literacy skills in the 21st-century education era can diversify learning activities. This is in line with Ng et al. (2023) research stating that technology is becoming increasingly popular for facilitating online learning for students. Based on these findings, the digital literacy skills of prospective teacher students need to be enhanced because aside from diversifying learning, digital literacy is also crucial for participating in the modern era to utilize the benefits of digital resources and information safely, securely, and sustainably.

Based on these needs, to enhance the digital literacy skills of prospective teacher students, a comprehensive and credible assessment is required to measure these skills. The assessment should be objective, thus necessitating high-quality tools or instruments to assist during the assessment process (Istiyono, 2020; Nurhayati, 2016; Rahman & Nasryah, 2019). Quality and credible assessment instruments for measuring digital literacy skills among prospective teacher students are still scarce and underdeveloped. These instruments should have proven validity, reliability estimation, difficulty level, and distractor properties (Arifin, 2017; IEA, 2020). Given this situation, the researcher is interested in developing a quality assessment instrument for digital literacy skills among prospective teacher students through validation, reliability estimation, difficulty level, and distractor properties.

Research related to digital literacy has been conducted in various contexts. For instance, Nur (2019) focused on the digital literacy skills of religious organization activists in Madrasah Aliyah Negeri (MAN) in Bandung, while Febliza and Okatariani (2020) aimed at developing digital literacy instruments for schools, teachers, and students. Both studies align thematically and conceptually with the current research, particularly regarding the development of assessment instruments to measure digital literacy. However, key differences exist, especially in the focus of the research subjects, instrument development procedures, and the evaluation of instrument quality. Nur's (2019) study emphasized activists in religious organizations, and Febliza and Okatariani (2020) targeted schools, teachers, and students. In contrast, this study focuses on prospective teacher students. Moreover, the instrument developed in this study follows Mardapi's (2017) development procedure and undergoes rigorous validation, reliability estimation, difficulty level, and distractor analysis. This research will ultimately produce a tested and validated instrument, providing a comprehensive profile of the digital literacy skills of prospective teacher students at state universities in Yogyakarta.

The instrument developed in this study is a practical self-assessment instrument. Self-assessment instruments can be used to measure digital literacy skills, indicating that digital literacy involves not only an individual's ability to master technology but also various other literacies that must be acquired, such as information literacy, visual literacy, media literacy, and communication literacy. Self-assessment instruments are suitable for digital literacy assessment based on the assumption that individuals know themselves best (Hardini, 2020; Laanpare et al., 2019). This instrument comprises 35 statements that assess seven aspects of digital literacy among prospective teacher students. The statements were organized and incorporated into a Google Form to facilitate online data collection and completion. Google Forms provide flexibility in data management and enable efficient further analysis.

The use of self-assessment instruments in this context aims to guide prospective teacher students to reflect on their experiences, knowledge, and skills in digital literacy. Additionally, the self-assessment instrument to be developed can also help them identify their weaknesses and motivate them to improve their digital literacy skills. This instrument will be tested on

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v28i2.72712

prospective teacher students in several education programs at state universities in the Special Region of Yogyakarta. Thus, this research aims to develop a digital literacy assessment instrument for prospective teacher students that can truly assess the extent of their digital literacy skills.

RESEARCH METHOD

This study aims to develop an assessment instrument for digital literacy skills among prospective teacher students. The developed assessment instrument for digital literacy skills is a self-assessment instrument using a Likert scale. Self-assessment is a method where participants independently examine or evaluate the extent to which they have achieved a characteristic or skill (R. Arifin et al., 2018; Hardini, 2020; Susilowati & Estidarsani, 2017). The response options in this instrument are divided into four categories: highly suitable (HS), suitable (S), not suitable (NS), and highly unsuitable (HU). The development of the assessment instrument in this study follows the 10-stage instrument development model, according to Mardapi (2017).

The subjects of this study are prospective teacher students at state universities in the Special Region of Yogyakarta. One of the courses taken by every prospective teacher-student is the micro-teaching course. The micro-teaching course trains students to prepare them for actual teaching (Prayitno et al., 2019). Therefore, the subjects are students who have taken the micro-teaching course.

Sample selection for the pilot test uses a purposive sampling technique. Purposive sampling allows for the determination of the characteristics of research subjects and their placement (Etikan et al., 2016). Purposive sampling is used to determine the criteria for research pilot test subjects. This limited-scale pilot test is conducted with ten prospective teacher students from the engineering faculty of Universitas Negeri Yogyakarta.

The large-scale pilot test uses the developed instrument revised based on expert validation results and limited-scale pilot testing. The subjects of this large-scale pilot test consist of 262 students who have taken the micro-teaching course from Universitas Negeri Yogyakarta and Universitas Islam Negeri Sunan Kalijaga Yogyakarta.

Data in this study consist of qualitative and quantitative data. Qualitative data comes from expert validation of content validity. Data sourced from experts are collected using instrument assessment sheets. The qualitative data collected include corrections, suggestions, and improvements to the developed instrument.

Quantitative data in this study are obtained from instrument assessment sheets by experts collected using Aiken's V with a rating scale from 1 (highly unsuitable) to 4 (highly suitable). The data generated are scores of item validity based on an examination of the suitability between statement items and other instrument components, such as aspects of digital literacy among prospective teacher students, developed indicators, item questions, and scoring guidelines. The results of this quantitative analysis serve as a reference for improvements before conducting the large-scale pilot test.

Content validity is performed by examining expert assessments of the developed instrument. The experts involved in this validation are three experts in educational assessment and three experts in educational technology. Content validation includes proving the face validity and logical validity of the instrument prepared in the initial instrument assembly stage. Construct validation is conducted based on the results of the large-scale pilot test using confirmatory factor analysis (CFA) and RStudio software to determine the construct validity of the instrument. The reliability estimation of the instrument will be analyzed using the Alpha-Cronbach reliability coefficient. The Alpha-Cronbach reliability coefficient in reliability estimation is used to estimate the reliability of multidimensional instruments, thus applicable to this instrument. Interpretation of the reliability coefficient calculation results can be done in

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v28i2.72712

various ways. One way to interpret the results of reliability estimation calculation is by using reliability coefficient categories. Reliability coefficient category by Istiyono (2020).

Item analysis is conducted during the development of this instrument to determine the quality of each instrument item. This item analysis uses item response theory to data from the large-scale pilot test. The type of item response theory analysis used is the graded response model (GRM).

The estimation of digital literacy measurement among prospective teacher students uses data from the large-scale pilot test. The total scores from the large-scale pilot test are then compared with digital literacy skill categories (Mardapi, 2012). The calculation of skill categories is done by converting instrument measurement results into numerical values based on the norm-referenced assessment system (PAN) (Istiyono, 2020).

FINDINGS AND DISCUSSION

The development study of this digital literacy assessment instrument begins with a review of documents related to digital literacy based on the assessment guidelines and evaluation of the national literacy movement by the Ministry of Education and Culture (Kemendikbud, 2017), as well as theories by several experts. Based on the documentation study and theories from these experts, seven aspects of digital literacy are used in this development study. These seven aspects include information literacy, digital scholarship, learning skills, ICT literacy, career and identity management, communication and collaboration, and media literacy (Ayu, 2021; Liansari & Nuroh, 2018; Yanti et al., 2021; Yunita & Indrawati, 2021). Next, conceptual definitions and operational definitions are formulated as references to determine the construction of the developed instrument.

These seven aspects of the instrument are then reviewed based on existing theories, resulting in a total of 16 indicators. These indicators are further developed into 40 statement items overall, which are considered capable of measuring each aspect in the assessment instrument for digital literacy among prospective teacher students. The next development stage involves proving content validity through expert judgment. The 40 statement items developed based on indicators derived from digital literacy among prospective teacher students are then given to three experts in educational assessment and three experts in educational technology. Validation sheets are provided to the experts to request assessments of the instrument's items that have been developed.

The experts provide assessments based on their expertise on the written items, where the assessments given consist of quantitative and qualitative evaluations. Quantitative assessments are given in the form of graded scores from 1 to 4. A score of 1 is given if the item is highly unsuitable for the digital literacy indicator of the aspect. In contrast, a score of 4 is given if the item is highly suitable for the digital literacy indicator of the aspect. Subsequently, the obtained scores (Appendix) are processed using Aiken's V formula to determine the content validity index for each item.

Based on the results of Aiken's V, it is evident that quantitatively, 38 out of 40 items are above the validity threshold of 0.78. For the two items that are below the validity threshold, specifically items 16 and 28, revisions were made to these two items. In addition to quantitative judgment, experts also provided qualitative input by giving suggestions and improvements to the developed instrument. One of the suggestions given was to add items from each indicator to make it three to five items. The suggestions regarding the writing of items were then used as a basis for refining the statement items in terms of grammar, context, and content to ensure understanding. Some items that have been improved include items 1, 6, 23, 24, 25, 28, 29, 31, 33, 36, 38, 39, and 40. After going through the improvement stage, the instrument was

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v28i2.72712

rearranged into a set of instruments comprising 48 items and submitted again to the experts for re-evaluation, which was then tested on prospective teacher students.

Next, a limited trial was conducted. Prospective teacher students involved in this limited-scale trial were ten students from Universitas Negeri Yogyakarta's mechanical engineering education program. The results of this limited-scale trial will serve as a reference for improving the instrument. Based on the review, it was found that the students encountered no difficulties while completing the instrument. Completing the 48-item instrument by students took a maximum of 30 minutes. Based on the results of this limited-scale trial, none of the instrument items needed further improvement or elimination.

Construct validity in the development of this instrument was assessed using data from the large-scale trial with confirmatory factor analysis (CFA). Confirmatory factor analysis (CFA) was performed to test the model's suitability based on the collected data and existing theories. Based on the CFA testing results before modification, it was found that the factor loading values for each item could be determined. Factor loading in CFA refers to the correlation coefficient between the measurement indicator variables (items) and the proposed factor or aspect. Factor loading is the numerical representation of the strength of the relationship between each item and the measured factor used to evaluate the construct validity of the model. Factor loading values that meet the criteria are those with factor loadings greater than 0.4 (Hair et al., 2019). Based on the CFA analysis results regarding factor loading values, it was found that there were 13 items with factor loading values below 0.4.

The next step involved testing model fit. Based on the model fit testing results, it was found that two fit index criteria were met, but most of the fit index criteria were not met, such as chi-square, goodness of fit index (GFI), normed fit index (NFI), incremental fit index (IFI), comparative fit index (CFI), Tucker-Lewis index (TLI), and adjusted goodness fit of index (AGFI). This indicates that the tested model does not match the empirical data or the constructed items are not valid for measuring the aspects of digital literacy among prospective teacher students.

Subsequently, modifications were made to the model to obtain a better model. Modifications in this research were made by evaluating the initial model based on factor loading and modification indices. Modifications based on the evaluation of the initial model using factor loading were done by removing or eliminating items with factor loadings below 0.4. The 13 items with factor loadings below 0.4 were eliminated. Modifications using indices in this research were done by correlating relationships between items or aspects. The researcher correlated relationships between items by considering substantial and theoretical connections between items or aspects. After going through the modification process, 35 items remained, which were then reanalyzed with CFA. Below is the factor loading table for the 35 items remaining after modification. Based on the CFA testing results after modification, it can be seen that the factor loading values for each item fall into the valid category with no factor loading values below 0.4 or $\lambda > 0.4$. Next, to assess whether the model fits in the CFA analysis, we look at the predetermined parameters.

In the model fit testing stage after modification, it was found that 4 out of 9 fit index criteria were still not met, namely chi-square, goodness of fit index (GFI), normed fit index (NFI), and adjusted Goodness Fit of Index (AGFI). Therefore, the root mean square error of approximation (RMSEA) value along with the Comparative Fit Index (CFI) and Tucker-Lewis Index (TLI) values can be an alternative to determine whether the model can be considered fit or not (Umar & Nisa, 2020). An RMSEA value below 0.05 with CFI and TLI values above 0.9 is sufficient to declare a model fit (Umar & Nisa, 2020). Based on this, the model in this study is considered fit because the RMSEA value is below 0.05, specifically 0.049; with CFI and TLI values above 0.9, specifically 0.915 and 0.908, respectively. Therefore, it can be concluded that the construct of the assessment instrument for digital literacy skills of prospective teacher

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v28i2.72712

students, consisting of 7 aspects, has been proven valid and significant in measuring latent variables (digital literacy skills). Additionally, the 35 items in the instrument have also been proven valid and significant for each of their respective aspects.

The next step is to estimate the reliability of the instrument. The reliability estimation in this study also utilizes RStudio software. As a reference, the categories of Cronbach's Alpha coefficient can be seen in the following table (Istiyono, 2020). The Cronbach's Alpha reliability coefficient is applied to measure the reliability of instruments with multiple dimensions, making this method suitable for use in the context of this instrument. A snapshot of the reliability estimation results using RStudio software based on Cronbach's Alpha reliability coefficient in this study can be seen in Figure 1.

```
Reliability analysis
Call: alpha(x = data)

  raw_alpha std.alpha G6(smc) average_r S/N    ase mean   sd median_r
       0.94      0.94    0.96      0.32  17 0.0052  3.2 0.43     0.32
```

Figure 1. Snapshot of Reliability Test Results with Alpha-Cronbach

Based on the excerpt above, the Alpha-Cronbach reliability coefficient of the instrument is 0.94. According to the reference, Alpha-Cronbach reliability coefficients ranging from 0.80 to 1.00 fall into the category of instruments with a very high level of estimated reliability or very reliable. Therefore, based on this, the assessment instrument for digital literacy skills of prospective teacher students in this study can be considered highly reliable.

Unidimensionality, a prerequisite for item response theory analysis, was rigorously met in this study. The data was meticulously analyzed using RStudio to generate a scree plot and total variance explained. The results of this comprehensive analysis can be seen in Figure 2.

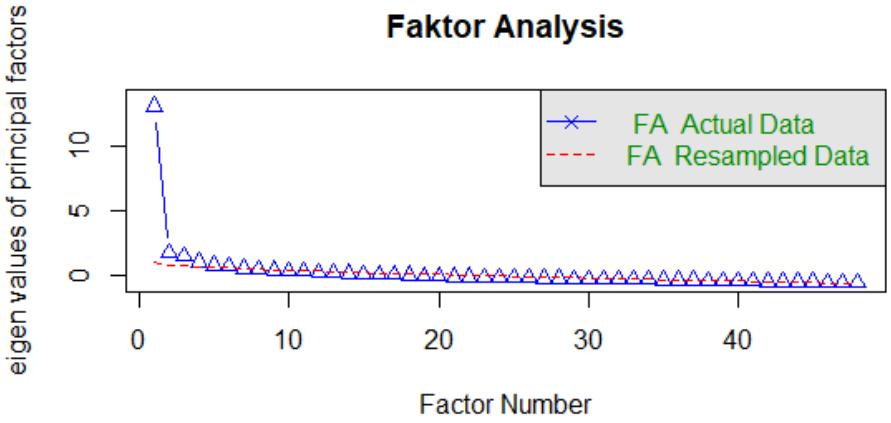

Figure 2. Scree Plot Analisis Faktor

Based on the illustration (Figure 2), the scree plot indicates that the main factor has a much steeper slope compared to the other factors. This suggests that the main extracted factor is significantly higher than the other factors, indicating that the test is unidimensional. This implies that the developed instrument measures a specific dimension, namely the digital literacy skills of prospective teacher students.

Local independence in item response theory reflects that each item in the instrument is answered independently. The term "independent" applies not only to the influence of other participants but also to the influence of the items in the instrument itself. The assumption of

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v28i2.72712

local independence is considered met when a participant's response to one item does not affect their response to other items. In this context, local independence is considered met because participants' response data have met the assumption of unidimensionality in the previous test. Therefore, it can be concluded that the assumption of local independence has been fulfilled.

Item parameter invariance in this study is implemented by dividing respondents into two groups, namely odd and even groups. The analysis method uses the Graded Response Model (GRM) with the assistance of RStudio software, and the results are displayed in Figure 3.

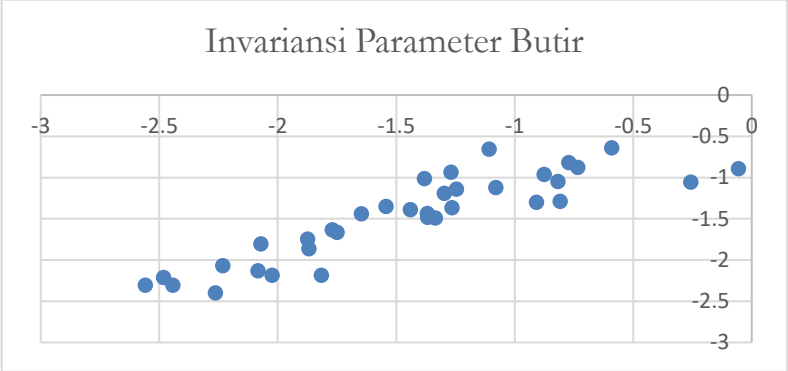

Figure 3. Item Parameter Invariance

Based on the scatterplot above, it can be seen that the distribution of points tends to approach the line x = y. Each point is relatively close to the diagonal line, indicating a similarity in the difficulty level of items between the two different groups. The implication is that there is no significant difference in the difficulty level of items tested between the two groups. Therefore, this result indicates that there is no variation in parameters between the odd and even groups. The analysis of this ability invariance is obtained through the application of the Multiple-Group GRM IRT model, which is specifically used to compare students' abilities in the odd and even groups. RStudio software is used as an analytical tool, and it will produce output in the form of a scatterplot that provides an overview of students' ability invariance.

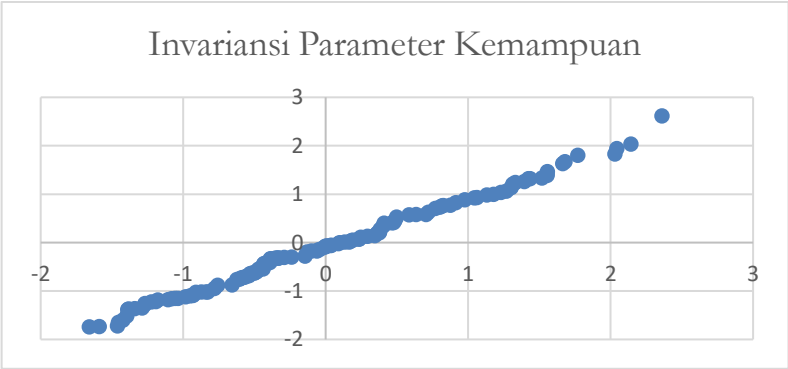

Figure 4. Skills Parameter Invariance

Based on the scatterplot above, it can be observed that the distribution of points tends to follow the line x = y. Each point is relatively close to the slope line, indicating that students' abilities remain consistent even when faced with items of varying difficulty levels and discrimination indices. This indicates the absence of parameter variation between odd and even groups.

The next analysis step is the item characteristic analysis using the IRT GRM model approach. The item characteristic analysis is conducted on response data to an instrument that has been proven valid in terms of content and construct. The instrument consists of 35 items

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v28i2.72712

that make up seven aspects. The analysis is carried out using RStudio software. At this stage, several results are obtained. First, the proposed factor model aligns with empirically obtained data based on the RMSEA value for 35 items, which is less than or equal to 0.05. Second, based on the analysis results, it is known which items have the highest discrimination (a) and difficulty level (bi) for each item. The item with the highest discrimination is item H43 with a = 2.837, and the item with the lowest discrimination is item G37 with a = 0.974. As for the difficulty level, item G37 has the lowest difficulty level, with $b1$, $b2$, and $b3$ being -4.975, -3.116, and 0.018, respectively. Third, it can be determined which responses are more likely to be chosen by student teachers based on their ability range. For example, for item G37, students with abilities less than -4.975 ($b1$ = -4.975) are likely to choose category 1 (very unsuitable), students with abilities between -4.975 and -3.116 ($b2$ = -3.116) are likely to choose category 2 (unsuitable), students with abilities between -3.116 and 0.018 ($b3$ = 0.018) are likely to choose category 3 (suitable), and students with abilities above 0.018 are likely to choose category 4 (very suitable). Fourth, it is known that the maximum value of the instrument's information function is obtained at an ability (θ) of -1.4 and SEM of 0.18. This means that the test provides optimal information when used with students with an ability of -1.4. Furthermore, the information function curve and SEM intersect at θ = -5.4 and 2.8, indicating that the instrument as a whole is suitable for testing students with abilities between -5.4 and 2.8. This is because the information function value reaches its peak, and the standard error value is at its minimum, approaching 0.

Next, the digital literacy skills of student teachers are based on their responses to the given instrument and processed with Maximum Likelihood Estimation. The results are analyzed using RStudio software, and the complete output can be seen in the appendix. The ability results obtained from MLE estimation are in logits ranging from -1.740 to 2.617. These scores are then converted into numerical values using the PAN system. The measurement of student teachers' skills aims to provide an overview of the digital literacy skills of student teachers at state universities in the Special Region of Yogyakarta. The results of the digital literacy skills measurement for student teachers are categorized into five categories: excellent, good, sufficient, poor, and very poor. The measurement results are presented in Table 1.

Table 1. Results of Digital Literacy Ability Measurement

No.	Kategori	Interval	Jumlah	Persentase (%)
1	Very Good	θ > 1.367	24	9%
2	Good	0.432<θ≤1.367	50	19%
3	Sufficient	-0.503< θ≤0.432	99	38%
4	Poor	-1.439< θ≤-0.503	79	30%
5	Very Poor	θ≤-1.439	10	4%

Based on the results of the digital literacy skills measurement, the proportion of each category is known. The very good category has a proportion of 9%, the good category has a proportion of 19%, the sufficient category has a proportion of 38%, the poor category has a proportion of 30%, and the category very poor has a proportion of 4%.

Based on the results of this study, it is known that the construction of the instrument was developed based on seven aspects, namely information literacy, digital scholarship, learning skills, ICT literacy, career and identity management, communication and collaboration, and media literacy. This aligns with previous experts' explanations, such as Bayrakci & Narmanlioğlu (2021), who stated that the aspects of digital literacy include ethics and responsibility, general knowledge and functional skills, daily usage, advanced production, privacy and security, and

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v28i2.72712

social dimension. The opinion of Bayrakci and Narmanlioğlu is similar to that of Komlayut and Srivatanakul (2017), who stated that digital literacy consists of six aspects: photo-visual digital skill reproduction, reproduction digital skill, branching digital skill, information digital skill, socio-emotional digital skill, and real-time digital skill. Meanwhile, Perdana et al. (2019) explained that the aspects of digital literacy are knowledge assembly, content evaluation, internet searching, and hypertextual navigation. These expert opinions, when elaborated with this research, are actually interrelated; however, this study emphasizes broader and deeper skills in the context of modern digital education. This is clearly seen in the aspects of digital scholarship and learning skills present in this study.

Furthermore, this study also produced an instrument that was tested through content validity using Aiken's V, construct validity using CFA, and item characteristics analysis through the IRT model GRM. These tests are similar to the instrument development conducted by Bashooir and Supahar (2018) and Lukitasari et al. (2022). These were carried out to ensure that the developed assessment instrument is valid, reliable, and capable of accurately and consistently measuring the intended constructs.

The final result of this study is a portrait of the measurement of digital literacy skills of prospective educator students at public universities in the Special Region of Yogyakarta. The measurement portrait shows that 9% are categorized as very good, 19% as good, 38% as fair, 30% as poor, and 4% as very poor. These measurement results can be used by prospective educator students at public universities in the Special Region of Yogyakarta to reflect and evaluate their learning related to digital literacy skills. Based on the items in the instrument, prospective educator students can undergo the necessary training or learning to improve their digital literacy skills. As Hardini (2020) stated, those who understand their situation best are themselves. Additionally, based on the measurement results from the instrument, certain institutions or organizations can organize seminars, outreach, or training needed by prospective educator students to enhance their digital literacy skills.

CONCLUSION

This development study yielded three main findings. (1) The construction of the instrument was based on seven aspects with 35 statement items: seven items for information literacy, six items for digital scholarship, three items for learning skills, three items for ICT literacy, seven items for career and identity management, four items for communication and collaboration, and five items for media literacy. (2) The quality of the developed instrument was tested through content validity using Aiken's V, construct validity with CFA, and item characteristics analysis using the GRM IRT model. (3) The profile of digital literacy skills measurement among prospective teacher students at state universities in the Special Region of Yogyakarta is as follows: 9% categorized as very good, 19% as good, 38% as sufficient, 30% as poor, and 4% as very poor.

Based on the findings of this research, it is recommended that the development of digital literacy assessment instruments for prospective teachers be continued by conducting further trials at private universities in the Special Region of Yogyakarta or other universities outside the Special Region of Yogyakarta that have different characteristics. These trials should be conducted under stricter supervision to ensure that students provide responses that reflect their actual experiences. Additionally, future researchers can develop other forms of assessment, such as peer assessment or behavioural observation. Peer assessment can provide additional perspectives on digital literacy skills and competencies that may not be revealed in self-assessment, while behavioural observation can provide objective data on how students use digital technology in real contexts. The combination of these assessment methods will provide a more comprehensive and in-depth understanding of students' digital literacy and enhance the validity and reliability of the assessment results. These recommendations for further research

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v28i2.72712

are expected to enrich our understanding of digital literacy and ways to enhance it, as well as ensure that the developed assessment instruments can be applied broadly and effectively. This additional research will also make an important contribution to developing the digital literacy competencies required in the current digital era.

Conflict of interests

There are no known conflicts of interest associated with this publication.

REFERENCES

Arifin, R., Kusumah, I. H., & Mubarak, I. (2018). Hasil Penilaian Diri Dan Penilaian Teman Sebaya Dibandingkan Dengan Assessment Dosen Untuk Hasil Produk Mata Kuliah Body Otomotif. Journal of Mechanical Engineering Education, 5(1), 79–80. https://doi.org/10.17509/jmee.v5i1.12623

Arifin, Z. (2017). Kriteria Instrumen dalam Suatu Penelitian. Jurnal THEOREMS (The Original Research of Mathematics), 2(1), 28–36.

Ayu, M. (2021). Implementasi Literasi Digital Pada Pembelajaran Fikih di MAN 1 Aceh Barat. UIN Syarif Hidayatullah Jakarta.

Bashooir, K., & Supahar. (2018). Validitas dan reliabilitas instrumen asesmen kinerja literasi sains pelajaran Fisika berbasis STEM. Jurnal Penelitian Dan Evaluasi Pendidikan, 22(2), 168–181. https://doi.org/10.21831/pep.v22i2.20270

Bayrakci, S., & Narmanlioğlu, H. (2021). Digital Literacy as Whole of Digital Competences: Scale development study. Düşünce ve Toplum Sosyal Bilimler Dergisi, 4, 1–30. www.dusuncevetoplum.org

Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of Convenience Sampling and Purposive Sampling Comparison of Convenience Sampling and Purposive Sampling. American Journal of Theoretical and Applied Statistics, 5 (1)(February), 1–4. https://doi.org/10.11648/j.ajtas.20160501.11

Falloon, G. (2020). From digital literacy to digital competence: the teacher digital competency (TDC) framework. Educational Technology Research and Development, 68(5), 2449. https://doi.org/10.1007/s11423-020-09767-4

Febliza, A., & Okatariani, O. (2020). Pengembangan Instrumen Literasi Digital Sekolah, Siswa Dan Guru. Jurnal Pendidikan Kimia Universitas Riau, 5(1), 1–9. https://doi.org/10.33578/jpk-unri.v5i1.7776

Fitriani, Yulis, P. A. R., Nurhuda, N., Sukarni, Gunita, A., & Cholijah, S. (2023). Peningkatan Literasi Digital Guru Dalam Pembelajaran Blended Learning. Jurnal Pengabdian UNDIKMA, 4(1), 140. https://doi.org/10.33394/jpu.v4i1.6536

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). Multivariate Data Analysis. https://doi.org/10.1002/9781119409137.ch4

Hardini, R. T. (2020). Pengembangan Instrumen Penilaian Pendidikan Karakter untuk Siswa Sekolah Dasar. In UNY. https://medium.com/@arifwicaksanaa/pengertian-use-case-a7e576e1b6bf

IEA. (2020). Reliability and Validity of International Large- Scale Assessment (H. Wagemaker (ed.)). Springer Nature Switzerland AG. https://doi.org/https://doi.org/10.1007/978-3-030-53081-5

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

Istiyono, E. (2020). Pengembangan Instrumen Penilaian dan Analisis Hasil Belajar Fisika (2nd ed.). UNY Press.

Kemendikbud. (2017). Pedoman Penilaian dan Evaluasi Gerakan Literasi Nasional (L. A. Mayani (ed.)). Kementrian Pendidikan dan Kebudayaan. http://gln.kemdikbud.go.id/glnsite/wp-content/uploads/2017/08/pedoman-penilaian-evaluasi-gln.pdf

Komlayut, S., & Srivatanakul, T. (2017). Assessing Digital Literacy Skills Using a Self-Administered Questionnaire. Review of Integrative Business and Economics Research, 6(3), 74–85. http://buscompress.com/journal-home.html

Laanpare, M., Senior Researcher, Tallinn Unuversity, & Centre for Educational Technology (2019). Recommendations on Assessment Tools for Monitoring Digital Literacy within UNESCO ' s Digital Literacy Global Framework (Issue 56). UNESCO Institute for Statistics. https://uis.unesco.org/sites/default/files/documents/ip56-recommendations-assessment-tools-digital-literacy-2019-en.pdf

Liansari, V., & Nuroh, E. Z. (2018). Realitas Penerapan Literasi Digital bagi Mahasiswa FKIP Universitas Muhammadiyah Sidoarjo. Proceedings of The ICECRS, 1(3), 250–251. https://doi.org/10.21070/picecrs.v1i3.1397

Lukitasari, M., Murtafiah, W., Ramdiah, S., Hasan, R., & Sukri, A. (2022). Constructing Digital Literacy Instrument and its Effect on College Students' Learning Outcomes. International Journal of Instruction, 15(2), 171–188. https://doi.org/10.29333/iji.2022.15210a

Mardapi, D. (2012). Pengukuran, Penilaian dan Evaluasi Pendidikan (1st ed.). Peter Lang Publishing.

Mardapi, D. (2017). Pengukuran Penilaian dan Evaluasi Pendidikan (2nd ed.). Parama Publishing.

Nasrullah, R., Aditya, W., Satya, T. I., Nento, M. N., Hanifah, N., Miftahussururi, & Akbari, Q. S. (2017). Gerakan Literasi Nasional. In L. A. Mayani (Ed.), Kementerian Pendidikan dan Kebudayaan. Kementerian Pendidikan dan Kebudayaan. https://gln.kemdikbud.go.id/glnsite/wp-content/uploads/2017/10/cover-materi-pendukung-literasi-finansial-gabung.pdf

Ng, D. T. K., Leung, J. K. L., Su, J., Ng, R. C. W., & Chu, S. K. W. (2023). Teachers' AI digital competencies and twenty-first century skills in the post-pandemic world. Educational Technology Research and Development, 71(1), 137–161. https://doi.org/10.1007/s11423-023-10203-6

Nur, M. (2019). Literasi Digital Keagamaan Aktivis Organisasi Keagamaan Di Madrasah Aliyah Negeri (Man) Di Kota Bandung. Jurnal SMART (Studi Masyarakat, Religi, Dan Tradisi), 5(1), 1–13. https://doi.org/10.18784/smart.v5i1.745

Nurhayati, A. (2016). Prinsip dan tujuan penilaian tindakan kelas. Jurnal Pembelajaran, 5(1), 7.

Nurlailah. (2022). Analisis Literasi Digital dan Keterampilan Berpikir Kritis Mahasiswa pada Pembelajaran Berbasis Web [Universitas Islam Negeri Raden Intan Lampung]. In UIN Raden Intan Lampung. https://doi.org/10.30998/rdje.v8i1.11738

Organisation for Economic Co-operation and Development (OECD). (2019). Programme for International Student Assessment (PISA) Result from PISA 2018. https://doi.org/10.1007/978-94-6209-497-0_69

ISSN (print) 2685-7111 | ISSN (online) 2338-6061

10.21831/pep.v28i2.72712

Perdana, R., Yani, R., Jumadi, J., & Rosana, D. (2019). Assessing Students' Digital Literacy Skill in Senior High School Yogyakarta. JPI (Jurnal Pendidikan Indonesia), 8(2), 169. https://doi.org/10.23887/jpi-undiksha.v8i2.17168

Prayitno, H. J., Sutama, Fuadi, D., Nurhidayat, Ratih, K., Harsono, Sufanti, M., Amalia, N., Sutarmi, S., Yuliana, I., & Hariyatmi. (2019). Desain dan Pedoman Pembelajaran Mikro (1st ed.). Muhammadiyah University Press.

Rahman, A. A., & Nasryah, C. E. (2019). Evaluasi Pembelajaran. In Uwais Inspirasi Indonesia (1st ed.). Uwais Inspirasi Indonesia. www.penerbituwais.com

Rahmawati. (2018). Analisis Kesenjangan Literasi Digital Mahasiswa Program Studi PAI PTKIN Indonesia [UIN Syarif Hidayatullah Jakarta]. In UIN Syarif Hidayatullah. http://repository.uinjkt.ac.id/dspace/handle/123456789/38121

Rukayah, Daryanto, J., Atmojo, I. R. W., Ardiansyah, R., Saputri, D. Y., & Salimi, M. (2022). Augmented Reality Media Development in STEAM Learning in Elementary Schools. Ingenierie Des Systemes d'Information, 27(3), 463. https://doi.org/10.18280/isi.270313

Sánchez-Cruzado, C., Santiago Campión, R., & Sánchez-Compaña, M. T. (2021). Teacher digital literacy: The indisputable challenge after covid-19. Sustainability (Switzerland), 13(4), 1–29. https://doi.org/10.3390/su13041858

Setyaningsih, R., Abdullah, A., Prihantoro, E., & Hustinawaty, H. (2019). Model Penguatan Literasi Digital Melalui Pemanfaatan E-Learning. Jurnal ASPIKOM, 3(6), 1202. https://doi.org/10.24329/aspikom.v3i6.333

Soeparno, P. L., & Ismaniati, C. (2022). Mathematics Pre-Service Teachers' Digital Literacy and Their Readiness Towards 21St Century Learning: a Mixed Method Study. BAREKENG: Jurnal Ilmu Matematika Dan Terapan, 16(2), 646–648. https://doi.org/10.30598/barekengvol16iss2pp643-650

Susilowati, & Estidarsani, N. (2017). Penerapan Self Assessment (Penilaian Diri) dengan Rubrik pada Hasil Belajar Siswa Menggambar Konstruksi Kusen Pintu Dan Jendela di Kelas X TGB SMKN 1 Kemlagi, Mojokerto. Jurnal Kajian Pendidikan Teknik Bangunan, 3, 243. tekniksipilunesa.org

Umar, J., & Nisa, Y. F. (2020). Uji Validitas Konstruk dengan CFA dan Pelaporannya. Jurnal Pengukuran Psikologi Dan Pendidikan Indonesia, 9(2), 7–8. https://doi.org/10.15408/jp3i.v9i2.16964

Wahyuningsih, S. (2021). Modul Literasi Digital Di Sekolah Dasar. Kementerian Pendidikan, Kebudayaan, Riset, Dan Teknologi, 1.

Warsiyah, Madrah, M., Muflihin, A., & Irfan, A. (2022). Urgensi Literasi Digital bagi Pendidik dalam Meningkatkan Keterampilan Mengelola Pembelajaran. Dimas: Jurnal Pemikiran Agama Untuk Pemberdayaan, 22(1), 124–126. https://doi.org/10.21580/dms.2022.221.10042

Yanti, N., Mulyati, Y., Sunendar, D., & Damaianti, V. (2021). Tingkat Literasi Digital Mahasiswa Indonesia. Diksa : Pendidikan Bahasa Dan Sastra Indonesia, 7(1), 61–63. https://doi.org/10.33369/diksa.v7i1.22391

Yunita, I., & Indrawati, I. (2021). Open Access jurnal dalam meningkatkan kemampuan literasi digital. Publishing Letters, 1(2), 12–13. https://doi.org/10.48078/publetters.v1i2.20

ISSN (print) 2685-7111 | ISSN (online) 2338-6061