

CONTENT VALIDATION OF A CRITICAL READING ASSESSMENT INSTRUMENT WITHIN A CRITICAL DIGITAL LITERACY FRAMEWORK USING THE FUZZY DELPHI METHOD

Veni Nurpadillah¹, Fathur Rokhman², Rahayu Pristiwati³

Faculty of Language and Arts, Universitas Negeri Semarang, Indonesia123

fathurrokhman@mail.unnes.ac.id

Abstract

This study aimed at developing and establishing content validity evidence for a multidimensional assessment instrument designed to assess students' critical reading skills in digital environment within the Critical Digital Literacy (CDL) framework. A research and development approach was employed with the Fuzzy Delphi Method (FDM) used to evaluate expert consensus regarding item relevance, clarity, and construct alignment. Ten experts in literacy, educational technology, and educational assessment participated in the content validation process. Responses were collected using a seven-point Likert scale and converted into Triangular Fuzzy Numbers (TFN). Item acceptance was determined using threshold values ($d \leq 0.20$) and defuzzification scores. The results indicated that all twenty-five items achieved expert consensus, with threshold values ranging from 0.113 to 0.182 and defuzzification scores ranging from 0.800 to 0.943. These findings provide evidence of strong content validity across five dimensions: critical digital literacy, critical reading, multimodality, linguistic clarity, and E-module usability. Rather than establishing complete instrument validity, the study offers an initial expert-based validation framework that can support subsequent empirical testing involving construct validity, reliability analysis, and field implementation.

Keywords: multimodal e-module, critical reading, critical digital literacy, Fuzzy Delphi, instrument validation

Introduction

Digital technology has transformed literacy from linear, text-based practices into multimodal, interactive, and dynamic forms of meaning-making (Mangen et al., 2019; Delgado et al., 2018). Reading in digital contexts needs comprehension, critical evaluation, synthesis, and reflection across visual, auditory, and interactive media (List & Alexander, 2019; Salmerón et al., 2020).

As digital information becomes more complex, students need to learn higher-order thinking skills like analysis, evaluation, and interpretation to deal with false information, fake news, and content driven by algorithms (Rouet & Britt, 2014; Bråten et al., 2020; Metzger & Flanagin,

2015; McGrew et al., 2018; Guess et al., 2020; Setyowati, 2020).

Although these studies demonstrate the growing importance of Critical Digital Literacy (CDL) in digital education, most existing research has primarily emphasized conceptual discussions or instructional interventions rather than the development of comprehensive assessment instruments. Consequently, there remains limited evidence regarding how CDL can be operationalized into measurable constructs capable of assessing students' critical reading performance across multimodal digital environments.

In this context, Critical Digital Literacy (CDL) offers a comprehensive framework for examining credibility, bias, power

dynamics, and ideological influences in digital texts (Hinrichsen & Coombs, 2014; Pangrazio et al., 2020; Nichols & LeBlanc, 2020). The rapid growth of digital information highlights the need for educational tools that promote critical and reflective engagement (Hatlevik et al., 2018).

Additionally, numerous digital literacy studies emphasize technical skills while neglecting the ideological and epistemological dimensions of digital texts (Hatlevik & Christophersen, 2013). Also, there are not many validated tools that combine Critical Digital Literacy, multimodality, and critical reading in a multidimensional way (Zumbo, 2007; Benson, 1998).

Existing assessment instruments predominantly focus on either digital literacy competencies, reading comprehension, or multimodal learning separately. Few instruments integrate these domains within a unified CDL framework, while empirical evidence regarding expert-based content validation of such multidimensional instruments remains limited. Consequently, there is still a need for an assessment framework that systematically combines critical reading, multimodal interpretation, and digital literacy competencies into coherent assessment constructs before proceeding to empirical validation.

The goals of this project are to develop and to test a multidimensional tool for measuring critical reading skills in digital settings, grounded in a CDL framework. The research utilizes the Fuzzy Delphi Method to quantify CDL into measurable dimensions, amalgamates multimodality and critical reading into a cohesive framework, and illustrates the efficacy of the Fuzzy Delphi Method in verifying intricate educational conceptions.

Literature Review

The progression of literacy in the digital age has transformed reading from linear,

text-centric methods to multimodal and critical interactions with information (Leu et al., 2017; Coiro, 2021). In digital contexts, meaning is constructed through textual, visual, auditory, and interactive elements (Kress, 2010; Jewitt, 2008). Critical Digital Literacy (CDL) has become a significant framework for analyzing the ideological, epistemological, and socio-political dimensions of digital texts (Pangrazio & Selwyn, 2021; Buckingham, 2015). CDL focuses on figuring out if something is credible, finding bias, and understanding how power works in digital information (Koltay, 2011; Livingstone, 2021), especially in relation to the rapid spread of fragmented, algorithmically generated content (van Deursen et al., 2021; Falloon, 2020).

Previous research on multimodal literacy has mainly focused on early education and basic meaning-making processes (Ranker & Janson, 2014; Grimminger et al., 2019), while studies on digital genres and higher-order critical analysis remain limited (Mills, 2016; Rowsell & Walsh, 2011). Although multimodal materials can improve engagement and comprehension (Morgan, 2013; Mayer, 2014), most research emphasizes accessibility over critical evaluation (Sung et al., 2016; Delgado et al., 2018). Digital literacy research often prioritizes technical skills over critical assessment of credibility, intent, and ideology (Hafidhi et al., 2024; van Laar et al., 2020; Wineburg & McGrew, 2019; McGrew et al., 2018).

There have been some improvements in CDL theory (Dooly & Darvin, 2022; Jiang & Gu, 2022), but it is still hard to make these ideas into useful evaluation tools (Spante et al., 2018; Reddy et al., 2020). Most current literacy research focuses on language skills and ignores the interconnected and multimodal aspects of digital literacy practices (Luke, 2012; Serafini, 2014). This study seeks to develop and evaluate a multidimensional instrument for assessing critical reading skills in digital

environments, grounded in a Critical Digital Literacy framework (Redecker, 2017; Carretero et al., 2017). The proposed instrument examines students' abilities to interpret, evaluate, and engage in critical thinking regarding digital news texts through various modalities.

Previous studies consistently acknowledge the importance of digital literacy and critical reading; however, they differ substantially regarding their conceptual emphasis. Some studies conceptualize digital literacy primarily as technological competence (Ng, 2012), whereas others argue that digital literacy should encompass

ideological awareness, critical evaluation, and sociocultural understanding (Pangrazio & Selwyn, 2021).

Likewise, multimodal literacy research emphasizes meaning construction across multiple semiotic modes (Kress, 2010), yet rarely incorporates assessment frameworks capable of measuring these competencies simultaneously. This conceptual fragmentation indicates that existing assessment approaches remain insufficient to evaluate critical reading comprehensively within digital environments.

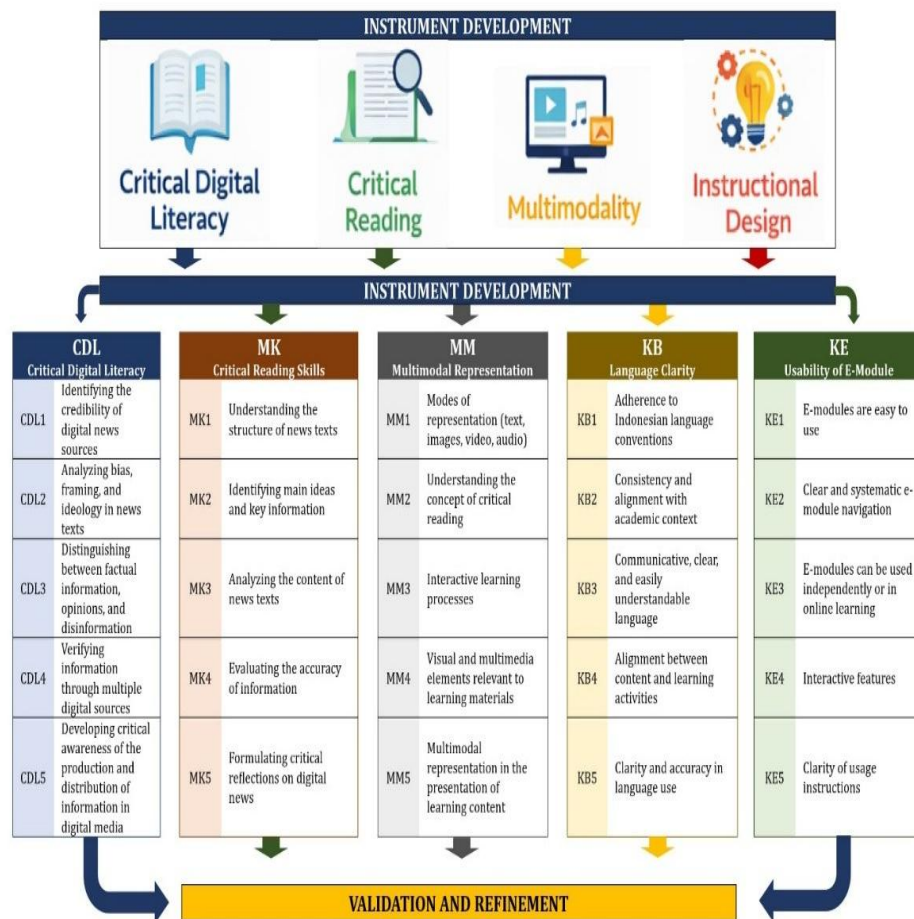


Figure 1. Conceptual framework of the multimodal e-module for digital news critical reading skills based on Critical Digital Literacy

Method

Research Design

This study employed a Research and Development (R&D) approach to develop an assessment instrument for evaluating

students' critical reading skills in digital environments within the Critical Digital Literacy (CDL) framework. According to Sukmadinata (2010, as cited in Putri et al., 2020), R&D integrates theoretical foundations with systematic empirical

procedures to develop educational products that are contextually relevant and scientifically grounded.

The present study focused specifically on the initial stage of instrument development, namely establishing content validity evidence through expert judgment. Rather than evaluating the instrument's psychometric performance, this study sought to determine whether the proposed assessment items adequately represented the intended theoretical constructs and were considered relevant, clear, and appropriate by domain experts. Empirical validation, including construct validity, reliability testing, and field implementation, remains beyond the scope of this study and is recommended for future research.

To establish expert-based content validity, the study employed the Fuzzy Delphi Method (FDM), an extension of the conventional Delphi technique that incorporates fuzzy set theory to accommodate ambiguity and uncertainty in expert judgments. FDM converts linguistic evaluations into mathematical representations, facilitating systematic consensus building while reducing subjectivity in expert decision-making. Previous studies have demonstrated that FDM is an effective approach for validating educational instruments, evaluating technologies, and developing instructional models (Anis & Masek, 2024). Within this study, FDM was used to determine expert consensus regarding the relevance, clarity, and construct alignment of each assessment item using threshold values and defuzzification analysis. It should be noted that the present study focuses exclusively on content validation through expert consensus. Empirical validation, including construct validity, reliability analysis, and field testing, is beyond the scope of this study and is recommended for future research..

Participants

The study involved 10 expert validators specializing in literacy, educational

technology, and educational assessment. The experts were purposively selected based on their academic qualifications and professional experience. They consisted of lecturers, researchers, and educational practitioners with expertise in digital literacy, multimodal learning, and educational assessment.

To ensure the credibility of the expert panel, participants were required to possess at least five years of professional experience, actively engage in research or scholarly publications, and demonstrate expertise in digital literacy, multimodality, or educational assessment. These criteria were established to ensure that expert judgments provided credible evidence for assessing the content relevance and conceptual appropriateness of the proposed instrument.

Research Procedures

1. Instrument development

The study began with the development of assessment items grounded in the theoretical foundations of Critical Digital Literacy, critical reading, and multimodal learning. A comprehensive literature review was conducted to identify the principal constructs and dimensions underlying the instrument, thereby ensuring alignment between theoretical concepts and assessment indicators.

2. Expert-based content validation

Following the initial development process, the instrument underwent expert-based content validation involving ten specialists in digital literacy, educational technology, and educational assessment. The experts evaluated each item with respect to its clarity, relevance, representativeness, and alignment with the intended constructs while also providing qualitative suggestions for refinement. This procedure was designed to establish evidence of content validity rather than empirical measurement validity.

3. Data collection

Expert judgments were collected using a 7-point Likert scale, allowing validators to indicate their level of agreement regarding each item's relevance and construct representation. The responses were subsequently transformed into Triangular Fuzzy Numbers (TFN) to facilitate Fuzzy Delphi analysis and accommodate uncertainty inherent in linguistic judgments.

4. Fuzzy delphi analysis

The collected data were analyzed using the Fuzzy Delphi Method. Expert ratings were converted into Triangular Fuzzy Numbers representing minimum, most plausible, and maximum values. Subsequently, threshold values (d) were calculated to determine the level of expert consensus, with $d \leq 0.20$ indicating acceptable agreement. Defuzzification analysis was then performed to obtain crisp values representing the relative importance and acceptability of each assessment item. Items satisfying the predefined consensus criteria were retained as evidence of content validity.

Data Analysis Techniques

This study employed the Fuzzy Delphi Method (FDM) as a systematic approach for establishing expert-based content validity. Initially, responses obtained from the seven-point Likert scale were transformed into Triangular Fuzzy Numbers (TFN) consisting of lower (l), middle (m), and upper (u) values to represent linguistic judgments quantitatively.

The threshold value (d) was subsequently calculated to determine the degree of agreement among experts. Items with $d \leq 0.20$ were considered to have achieved acceptable consensus, indicating that experts consistently judged the items as relevant to the intended constructs. A defuzzification procedure was then conducted to generate crisp values reflecting the perceived importance and acceptability of each assessment item.

Finally, decisions regarding item retention were based on the combined results of threshold values and defuzzification scores. Items meeting the predetermined consensus criteria were retained as evidence supporting their content validity, whereas items failing to meet these criteria would require revision or elimination. Accordingly, the present analysis provides evidence regarding expert consensus on item quality rather than empirical evidence of the instrument's psychometric properties, which should be established through subsequent studies involving reliability analysis, construct validation, and field testing.

Finding and Discussion

The Fuzzy Delphi analysis indicated a high level of expert consensus across all 25 proposed assessment items. Threshold values (d) ranged from 0.113 to 0.182, all of which were below the recommended consensus criterion of 0.20. These findings indicate that the participating experts consistently agreed that the assessment items were relevant, clearly formulated, and appropriately represented the intended theoretical constructs. Consequently, all items satisfied the predefined criteria for content validity and were retained for subsequent stages of instrument development.

As illustrated in Figure 2, an inverse relationship was observed between threshold values and defuzzification scores. Assessment items with lower threshold values generally obtained higher defuzzification scores, indicating stronger agreement among experts regarding their relevance and representativeness.

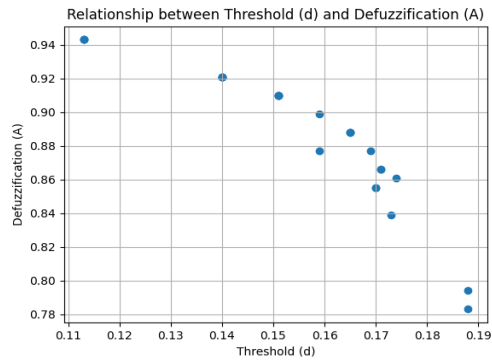


Figure 2. Relationship between threshold values (d) and defuzzification values (A)

This pattern suggests a consistent level of expert consensus across the proposed assessment items. Rather than demonstrating the instrument's overall psychometric quality, these findings provide preliminary evidence supporting the content validity of the assessment framework and justify its progression to subsequent empirical validation.

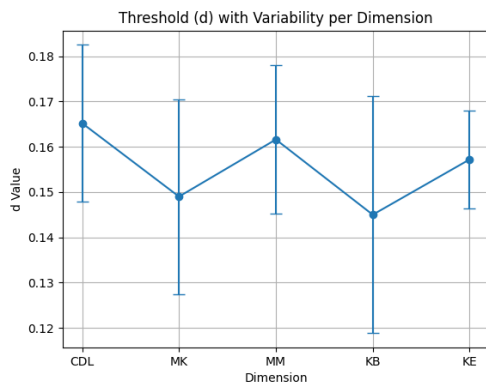


Figure 3. Threshold Values Across Instrument Dimensions

Figure 3 demonstrates that all instrument dimensions achieved threshold values below the recommended criterion of 0.20, indicating consistent agreement among experts across all proposed constructs. Items such as MK3, MK4, MM5, and KB3–KB5 obtained the lowest threshold values, reflecting particularly strong expert consensus regarding their relevance. These findings suggest that the indicators representing analytical evaluation, multimodal interpretation, and linguistic clarity were consistently perceived by

experts as important components of the proposed assessment framework.

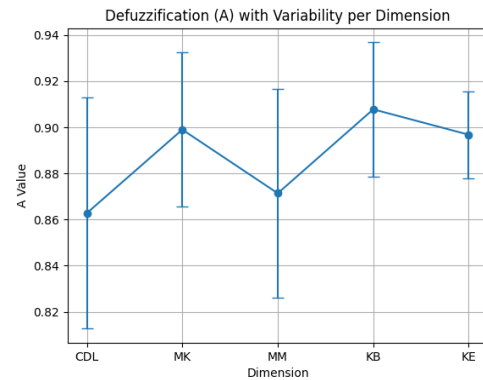


Figure 4. Defuzzification Values

Figure 4 shows that all assessment items achieved defuzzification scores exceeding 0.800, indicating a high level of expert acceptance. The highest scores were observed for indicators associated with analytical evaluation and critical interpretation, suggesting that experts considered these competencies central to assessing critical reading within digital environments. The consistently high defuzzification scores provide further evidence supporting the content relevance of the proposed assessment items.

Figure 4 illustrates that all defuzzification values (A) above 0.800, indicating a consistently high level of expert consensus and endorsement across all metrics. The strongest results focused on analytical and evaluative skills, underscoring the need for higher-order thinking in digital critical reading, including assessing credibility, recognizing assumptions, and synthesizing meaning from multimodal sources. These findings underscore the conceptual and empirical importance of analytical and evaluative aspects in evaluating critical reading ability in intricate digital contexts.

Meanwhile, the Multimodality dimension highlights the integration of text, images, animations, and interactive media, requiring learners to construct meaning across multiple modes. The aspects of

Linguistic Clarity and E-module Usability enhance accessibility, readability, and user-centered design, while diminishing

cognitive load and augmenting engagement and understanding.

Code	Item Description	Validator(s)										Average TFN			Defuzzification Threshold					Decision
		V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	L_avg	m_avg	u_avg	A	l	m	u	d	
CDL 1	Identifying the credibility of digital news sources	5	6	6	6	6	5	6	6	5	6	0,619	0,782	0,949	0,783	0,071	0,163	0,330	0,188	✓ Accepted
CDL 2	Analyzing bias, framing, and ideology in news texts	5	6	7	6	6	6	7	6	6	6	0,685	0,848	0,983	0,839	0,058	0,163	0,298	0,173	✓ Accepted
CDL 3	Distinguishing between factual information, opinions, and disinformation	7	7	7	6	7	6	7	7	6	7	0,782	0,949	1,000	0,910	0,067	0,167	0,218	0,151	✓ Accepted
CDL 4	Verifying information through multiple digital sources	7	7	7	6	7	7	6	7	7	7	0,798	0,966	1,000	0,921	0,051	0,168	0,202	0,140	✓ Accepted
CDL 5	Developing critical awareness of the production and distribution of information in digital media	6	7	6	6	7	7	5	6	6	7	0,717	0,882	0,983	0,861	0,090	0,165	0,266	0,174	✓ Accepted
MK 1	Understanding the structure of news texts	6	6	6	6	6	6	6	7	7	6	0,702	0,864	1,000	0,855	0,051	0,162	0,298	0,170	✓ Accepted
MK 2	Identifying main ideas and key information	6	6	7	6	6	6	6	7	7	6	0,718	0,881	1,000	0,866	0,067	0,163	0,282	0,171	✓ Accepted
MK 3	Analyzing the content of news texts	7	7	7	7	7	7	7	7	7	7	0,830	1,000	1,000	0,943	0,000	0,170	0,170	0,113	✓ Accepted
MK 4	Evaluating the accuracy of information	7	7	7	7	7	7	6	7	6	7	0,798	0,966	1,000	0,921	0,051	0,168	0,202	0,140	✓ Accepted
MK 5	Formulating critical reflections on digital news	6	7	6	7	7	7	7	7	6	7	0,782	0,949	1,000	0,910	0,067	0,167	0,218	0,151	✓ Accepted
MM 1	Modes of representation (text, images, video, audio)	6	6	6	7	5	6	6	6	6	5	0,635	0,799	0,949	0,794	0,081	0,170	0,314	0,188	✓ Accepted
MM 2	Understanding the concept of critical reading	6	6	7	7	6	6	6	6	6	6	0,702	0,864	1,000	0,855	0,051	0,162	0,298	0,170	✓ Accepted
MM 3	Interactive learning processes	7	7	7	7	6	7	7	6	7	6	0,782	0,949	1,000	0,910	0,067	0,167	0,218	0,151	✓ Accepted
MM 4	Visual and multimedia elements relevant to learning materials	6	6	6	6	6	7	7	7	7	6	0,734	0,898	1,000	0,877	0,077	0,164	0,266	0,169	✓ Accepted
MM 5	Multimodal representation in the presentation of learning content	7	7	7	7	7	7	7	6	6	7	0,798	0,966	1,000	0,921	0,051	0,168	0,202	0,140	✓ Accepted
KB 1	Adherence to Indonesian language conventions	6	6	6	7	6	7	7	7	7	6	0,750	0,915	1,000	0,888	0,080	0,165	0,250	0,165	✓ Accepted
KB 2	Consistency and alignment with academic context	6	7	7	7	6	7	6	7	6	6	0,750	0,915	1,000	0,888	0,080	0,165	0,250	0,165	✓ Accepted
KB 3	Communicative, clear, and easily understandable language	7	7	7	7	7	7	7	7	7	7	0,830	1,000	1,000	0,943	0,000	0,170	0,170	0,113	✓ Accepted
KB 4	Alignment between content and learning activities	6	7	6	6	7	6	6	7	6	7	0,734	0,898	1,000	0,877	0,077	0,164	0,266	0,169	✓ Accepted
KB 5	Clarity and accuracy in language use	7	7	7	7	7	7	7	7	7	7	0,830	1,000	1,000	0,943	0,000	0,170	0,170	0,113	✓ Accepted
KE 1	E-modules are easy to use	6	6	6	7	6	7	7	7	7	6	0,750	0,915	1,000	0,888	0,080	0,165	0,250	0,165	✓ Accepted
KE 2	Clear and systematic e-module navigation	6	6	7	7	6	6	6	7	6	6	0,718	0,881	1,000	0,866	0,067	0,163	0,282	0,171	✓ Accepted
KE 3	E-modules can be used independently or in online learning	6	6	7	6	7	7	7	7	6	7	0,766	0,932	1,000	0,899	0,077	0,166	0,234	0,159	✓ Accepted
KE 4	Interactive features	7	7	7	7	6	6	7	7	7	6	0,782	0,949	1,000	0,910	0,067	0,167	0,218	0,151	✓ Accepted
KE 5	Clarity of usage instructionsClarity of usage instructions	7	7	7	7	6	7	7	7	7	6	0,798	0,966	1,000	0,921	0,051	0,168	0,202	0,140	✓ Accepted

Figure 5. Summary of Fuzzy Delphi Analysis for Multidimensional Instrument Validation

The overall results presented in Figure 5 demonstrate that all proposed assessment items satisfied the three predefined Fuzzy Delphi criteria, namely acceptable threshold values ($d \leq 0.20$), high levels of expert consensus, and defuzzification scores above the recommended acceptance level. Collectively, these findings provide evidence supporting the content validity of the assessment items. Accordingly, all items were retained for future empirical evaluation. Nevertheless, these results should not be interpreted as evidence of the instrument's construct validity, reliability, or practical effectiveness, which require subsequent psychometric testing involving representative student samples.

Overall, the Fuzzy Delphi analysis provides expert-based evidence supporting the content validity of the proposed assessment instrument. The findings demonstrate that the assessment items are conceptually relevant, clearly formulated, and representative of the intended constructs according to expert judgment. However, because the present study focuses exclusively on expert consensus, further empirical studies involving pilot testing, construct validation, and reliability analysis are required before the instrument can be

considered fully validated for educational measurement.

Discussion

The findings of this study provide initial evidence supporting the content validity of the proposed assessment instrument for critical reading within the Critical Digital Literacy (CDL) framework. The high level of expert consensus indicates that the assessment items were considered relevant, representative, and conceptually aligned with the intended dimensions of Critical Digital Literacy, Critical Reading, Multimodality, Linguistic Clarity, and E-module Usability. Accordingly, the findings demonstrate that the proposed indicators adequately represent the theoretical constructs from the perspective of domain experts. However, these results should be interpreted as evidence of content validity rather than as confirmation of the instrument's psychometric performance, which requires subsequent empirical validation involving representative learners.

The proposed instrument conceptualizes critical reading as a multidimensional construct that integrates cognitive processing, multimodal meaning-making, and critical evaluation within digital

environments. This perspective is consistent with the view of Critical Digital Literacy as a contextual and sociocultural practice rather than merely a set of technical skills (Lankshear & Knobel, 2011; Street, 2003). By organizing these dimensions into measurable assessment indicators, the study contributes to the operationalization of Critical Digital Literacy for educational assessment while providing a theoretically grounded framework for future empirical investigation.

These findings also support contemporary perspectives that extend digital literacy beyond technological competence toward critical engagement with information. Previous studies emphasize that digital literacy involves evaluating credibility, recognizing bias, and interpreting meaning across multiple semiotic modes (Ng, 2012; Eshet, 2012; Dooly & Darvin, 2022; Jiang & Gu, 2022; Leu et al., 2015; Coiro, 2021). The present study contributes to this body of literature by demonstrating that experts consistently regarded these competencies as essential components of a critical reading assessment framework. Nevertheless, the study does not claim that these dimensions have been empirically confirmed as latent constructs because such evidence requires psychometric analyses using learner data.

The strong level of expert agreement further suggests that the proposed indicators possess substantial conceptual coherence and representativeness. Within the Fuzzy Delphi Method, consensus reflects convergence in expert judgments regarding the appropriateness of assessment items rather than empirical evidence of measurement quality (Ahmad et al., 2025). This interpretation is consistent with previous methodological studies demonstrating that FDM is an effective procedure for establishing expert-based content validity during the early stages of instrument development (Hsu, 2007; Okoli & Pawlowski, 2004; Hasson et al., 2000). Therefore, the present findings should be interpreted as providing

preliminary evidence supporting item quality before subsequent construct validation and reliability testing.

The integration of multimodality within the proposed framework reflects the evolving nature of reading practices in digital environments, where learners construct meaning through combinations of textual, visual, auditory, and interactive information. This interpretation is consistent with multimodal literacy theories proposed by Kress (2010) and Jewitt (2008), as well as studies demonstrating that multimodal learning environments can enhance learner engagement, comprehension, and interpretative flexibility (Mayer, 2014; Sung et al., 2016; Moreno & Mayer, 2007; Schnotz & Bannert, 2003). The present findings indicate that experts considered these multimodal components conceptually appropriate for inclusion within the proposed assessment framework. However, further empirical studies are required to determine whether these indicators function effectively when administered to students in authentic educational settings.

Methodologically, this study also illustrates the usefulness of the Fuzzy Delphi Method for establishing expert consensus during the initial stages of educational instrument development. Previous research has consistently shown that FDM provides a systematic and transparent approach for evaluating complex educational constructs by integrating expert judgments while reducing uncertainty in decision-making (Anis & Masek, 2024; Daud et al., 2024; Dalkey & Helmer, 1963; Skulmoski et al., 2007). In the present study, FDM successfully facilitated the identification of assessment items that experts considered relevant and representative of the intended constructs. Nevertheless, FDM should be viewed as providing evidence of content validity rather than comprehensive evidence of instrument validity.

Overall, this study contributes theoretically by demonstrating how Critical Digital

Literacy can be translated into measurable assessment indicators through a systematic expert-consensus process. Practically, the resulting assessment framework provides a foundation for future psychometric evaluation, including pilot testing, construct validation, reliability analysis, and large-scale implementation. Consequently, the present findings should be regarded as an important preliminary stage in instrument development rather than as evidence that the instrument has been fully validated for educational measurement.

Research Implications

The findings of this study have methodological, theoretical, and practical implications for the development of assessment instruments within the field of Critical Digital Literacy (CDL).

Methodologically, this study demonstrates that the Fuzzy Delphi Method (FDM) provides a systematic approach for establishing expert-based content validity during the early stages of instrument development. By integrating expert judgments through fuzzy set theory, FDM enables researchers to evaluate the relevance, clarity, and representativeness of assessment items while reducing subjectivity in decision-making. These findings are consistent with previous studies highlighting the effectiveness of FDM for educational instrument development and expert consensus building (Hsu, 2007; Okoli & Pawlowski, 2004; Hasson et al., 2000; Skulmoski et al., 2007; Linstone & Turoff, 2002).

Theoretically, this study contributes to the operationalization of the Critical Digital Literacy framework by translating its conceptual dimensions into measurable assessment indicators. Previous studies have emphasized the conceptual importance of digital literacy (Eshet, 2012; Ng, 2012; Pangrazio & Selwyn, 2021; Dooly & Darvin, 2022; Jiang & Gu, 2022; Spante et al., 2018), yet relatively few have

focused on developing assessment indicators capable of representing these multidimensional constructs. The present study therefore provides an initial conceptual framework that may support future psychometric investigations of Critical Digital Literacy assessment.

Practically, the proposed assessment framework offers a preliminary foundation for the development of assessment instruments intended to evaluate students' critical reading competencies in digital environments. However, the present findings should be interpreted as evidence of content validity derived from expert consensus rather than evidence of empirical measurement quality. Consequently, the instrument should undergo further empirical evaluation—including pilot testing, construct validation, reliability analysis, and field implementation—before being recommended for broader educational use (Redecker, 2017; Ferrari, 2013; Carretero et al., 2017; Ilomäki et al., 2016; Mayer, 2014; Clark & Mayer, 2016; Sung et al., 2016; Moreno & Mayer, 2007; Schindler et al., 2017; Lai & Bower, 2019).

Furthermore, this study provides a foundation for subsequent instrument validation research. Consistent with recommendations by Boateng et al. (2018), DeVellis (2016), Hair et al. (2019), and Taber (2018), future studies should establish additional psychometric evidence through reliability testing, exploratory and confirmatory factor analyses, and large-scale validation involving diverse educational contexts. Such investigations will strengthen the instrument's measurement properties and determine its applicability for educational assessment and future research.

Research Limitations

This study has several limitations. First, even if the number of validators met the Fuzzy Delphi Method's basic standards, the results would be more reliable and applicable to a wider range of situations if

the group included a greater variety of experts. The study primarily examines content validation using the Fuzzy Delphi Method, indicating that empirical validation has not yet been conducted. We need to conduct further studies to assess how well the tool performs in the real world, including testing its reliability and construct validity.

Another reason the study's results cannot be applied to other situations is that it was conducted at a specific type of school. We need to conduct more research to improve the instrument so it can be used in more situations and fields. This would make it more useful. The study predominantly employed numerical data, which inadequately represented all expert opinions. Subsequent research should utilize diverse methodologies to obtain more comprehensive and nuanced results.

Finally, a 7-point Likert scale may not fully capture the intricacies of expert evaluations, potentially rendering it less responsive to minor rating fluctuations. Future research could explore sophisticated scaling techniques to improve measurement accuracy. As emphasized in previous instrument validation studies, the caliber of an instrument is closely associated with its validity and reliability (Boateng et al., 2018).

Conclusions

This study provides expert-based evidence supporting the content validity of a multidimensional assessment instrument designed to assess critical reading within the Critical Digital Literacy (CDL) framework. Using the Fuzzy Delphi Method, all proposed assessment items achieved the predefined consensus criteria, indicating that experts considered the items relevant, representative, and conceptually aligned with the intended constructs. These findings support the retention of all assessment items for subsequent stages of instrument development.

The study also demonstrates that the Fuzzy Delphi Method is an effective approach for establishing content validity evidence during the initial phase of educational instrument development. Through systematic expert consensus, the study translated the conceptual dimensions of Critical Digital Literacy into measurable assessment indicators that provide a theoretically grounded framework for future assessment development.

However, the findings should be interpreted within the scope of content validation. The present study does not provide empirical evidence regarding the instrument's construct validity, reliability, or measurement performance. Consequently, further studies involving pilot testing, exploratory and confirmatory factor analyses, reliability testing, and field implementation are required before the instrument can be considered fully validated for educational assessment.

Overall, this study contributes to the growing body of research on Critical Digital Literacy by providing an initial assessment framework and expert-based content validity evidence that may serve as a foundation for subsequent psychometric validation and broader educational applications.

Suggestion

To enhance the empirical evaluation of the instrument, subsequent research should incorporate reliability testing, concept validation, and field deployment in various educational contexts. The instrument's validity, reliability, and overall efficacy in practical settings can only be determined through comprehensive trials with a heterogeneous population.

Furthermore, future research should employ mixed-methods approaches that combine quantitative metrics with qualitative data, such as expert assessments and user experiences. This amalgamation can enhance our comprehension of the instrument's functionality across various

contexts and improve the reliability and interpretability of the data.

Additionally, additional study is needed to create better evaluation tools, such adaptive rating systems or extended Likert scales, that can pick up on people's small shifts in mood and thought patterns. If the instrument is to be used, used, and scientifically legitimate in future research and teaching, it has to be more rigorous empirically, have more methodological variation, and have more precise measurements, according to these remarks.

References

- Ahmad, N. A., Hayaty, N., & Kasim, A. (2025). Defining the multifaceted roles of dental academics: A consensus approach using the Fuzzy Delphi Method Abstract: *PLOS One*, 8(August), 1–16.
- Anis, S. K., & Masek, A. (2024). Indonesian Journal of Educational Research and Technology: Fuzzy Delphi Method Application Towards Usability Evaluation of Problem-Solving Framework: Application in Education. *Indonesian Journal of Educational Research and Technology*, 4(1), 49–62.
- Benson, J. (1998). Developing a strong program of construct validation: A test anxiety example. *Educational Measurement: Issues and Practice*, 17(1), 10–17.
- Boateng, G. O., Neilands, T. B., & Frongillo, E. A. (2018). Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Frontiers in Public Health*, 6(June), 1–18.
<https://doi.org/10.3389/fpubh.2018.00149>
- Bråten, I., Strømsø, H. I., & Salmerón, L. (2020). Trust and mistrust when students read multiple information sources. *Educational Psychologist*, 55(1), 1–17.
- Buckingham, D. (2015). Defining Digital Literacy—What Do Young People Need to Know about Digital Media? *Nordic Journal of Digital Literacy*, 10, 21–35.
<https://doi.org/10.18261/ISSN1891-943X-2015-Jubileumsnummer-03>
- Carretero, S., Vuorikari, R., & Punie, Y. (2017). DigComp 2.1. Publications Office of the European Union.
- Coiro, J. (2021). Toward a multifaceted heuristic of digital reading. *Reading Research Quarterly*, 56(1), 9–31.
<https://doi.org/10.1002/rrq.302>
- Dalkey, N., & Helmer, O. (1963). An experimental application of the Delphi method. *Management Science*, 9(3), 458–467.
- Daud, A., Lina, S., Sitio, M., Bakri, A. A., & Judijanto, L. (2024). The Impact of the Fuzzy Delphi Technique on Technology Methods Based on Design Thinking. *Journal of Education Technology*, 8(1), 95–105.
- Delgado, P., Vargas, C., Ackerman, R., & Salmerón, L. (2018). Don't throw away your printed books: A meta-analysis on the effects of reading media on reading comprehension. *Educational Research Review*, 25, 23–38.
<https://doi.org/10.1016/j.edurev.2018.09.003>
- DeVellis, R. F. (2016). Scale development: Theory and applications (4th ed.). Sage.
- Dooly, M., & Darwin, R. (2022). Intercultural communicative competence in the digital age: Critical digital literacy and inquiry-based pedagogy. *Language and Intercultural Communication*, 22(3), 354–366.
- Eshet, Y. (2012). Thinking in the Digital Era : A Revised Model for Digital Literacy. *Issues in Informing Science and Information Technology*, 9(1), 267–276.
- Ferrari, A. (2013). DIGCOMP: A framework for developing digital

- competence in Europe. European Commission.
- Guess, A. M., Nagler, J., & Tucker, J. (2020). Less than you think: Prevalence of fake news. *Science Advances*, 6(14), eaay3539.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage.
- Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, 32(4), 1008–1015. <https://doi.org/10.1046/j.1365-2648.2000.t01-1-01567.x>
- Hatlevik, O. E., & Christophersen, K. A. (2013). Digital competence at the beginning of upper secondary school. *Computers & Education*, 63, 240–247.
- Hatlevik, O. E., Guðmundsdóttir, G. B., & Loi, M. (2018). Examining factors predicting students' digital competence. *Computers & Education*, 118, 1–13.
- Hill-May, A., et al. (2013). Student learning outcomes with multimedia. *Journal of Educational Multimedia and Hypermedia*.
- Hinrichsen, J., & Coombs, A. (2014). The five resources of critical digital literacy: A framework for curriculum integration. *Research in Learning Technology*, 21. <https://doi.org/10.3402/rlt.v21.21334>
- Hsu, C. (2007). The Delphi Technique: Making Sense Of Consensus. *Practical Assessment, Research & Evaluation*, 12(10), 1–8.
- Ilomäki, L., Paavola, S., Lakkala, M., & Kantosalo, A. (2016). Digital competence framework. *Educational Technology & Society*, 19(3), 655–669.
- Jewitt, C. (2008). Multimodality and literacy in school classrooms. *Review of Research in Education*, 32(1), 241–267. <https://doi.org/10.3102/0091732X07310586>
- Jiang, L., & Gu, M. M. (2022). Toward a Professional Development Model for Critical Digital Literacies in TESOL. *TESOL Quarterly*, 56(3).
- Koltay, T. (2011). The media and the literacies: Media literacy, information literacy, digital literacy. *Media, Culture & Society*, 33(2), 211–221. <https://doi.org/10.1177/0163443710393382>
- Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.
- Lai, J. W. M., & Bower, M. (2019). How is the use of technology in education evaluated? *Computers & Education*, 133, 27–42.
- Lankshear, C., & Knobel, M. (2011). *New literacies: Everyday practices and social learning* (3rd ed.). Open University Press.
- Leu, D. J., et al. (2015). The new literacies of online research and comprehension. *Educational Psychologist*, 50(1), 37–59. <https://doi.org/10.1080/00461520.2015.1020337>
- List, A., & Alexander, P. A. (2019). Toward an integrated framework of multiple text use. *Educational Psychologist*, 54(1), 20–38.
- Linstone, H. A., & Turoff, M. (2002). *The Delphi method: Techniques and applications*. Addison-Wesley.
- Livingstone, S. (2021). Digital literacy and the challenge of fake news. *European Journal of Communication*, 36(3), 1–15.
- Mangen, A., Walgermo, B. R., & Brønnick, K. (2019). Reading linear texts on paper vs screens. *International Journal of Educational Research*, 58, 61–68.
- Mayer, R. E. (2014). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd ed., pp. 43–71). Cambridge University Press.
- McGrew, S., Ortega, T., Breakstone, J., & Wineburg, S. (2018). The challenge

- that's bigger than fake news. *American Educator*, 42(3), 4–9.
- Metzger, M. J., & Flanagin, A. J. (2015). Psychological approaches to credibility assessment. *Journal of Communication*, 65(3), 445–466.
- Mills, K. A. (2016). Literacy theories for the digital age. *Social Semiotics*, 26(1), 1–9.
- Moreno, R. (2010). Cognitive load theory. *Educational Psychologist*.
- Morgan, H. (2013). Multimodal children's e-books help young learners in reading. *Early Childhood Education Journal*, 41(6), 477–483. <https://doi.org/10.1007/s10643-013-0575-8>
- Ng, W. (2012). Can we teach digital natives digital literacy? *Computers & Education*, 59(3), 1065–1078. <https://doi.org/10.1016/j.compedu.2012.04.016>
- Nichols, T. P., & LeBlanc, R. J. (2020). Beyond skills: Critical digital literacy. *Journal of Literacy Research*.
- Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool. *Information & Management*, 42(1), 15–29. <https://doi.org/10.1016/j.im.2003.11.002>
- Pangrazio, L., & Selwyn, N. (2021). Towards a school-based 'critical digital literacy' framework. *Learning, Media and Technology*, 46(1), 1–14. <https://doi.org/10.1080/17439884.2020.1771527>
- Redecker, C. (2017). European Framework for the Digital Competence of Educators: DigCompEdu. Publications Office of the European Union. https://publications.jrc.ec.europa.eu/repository/bitstream/JRC107466/pdf_digcomedu_a4_final.pdf
- Rouet, J.-F., & Britt, M. A. (2014). Relevance processes in multiple document comprehension. *Handbook of educational psychology*.
- Rowsell, J., & Walsh, M. (2011). Rethinking literacy education in new times. *Australian Journal of Language and Literacy*, 34(1), 53–62.
- Salmerón, L., García, A., & Vidal-Abarca, E. (2020). Reading multiple sources. *Educational Psychology Review*.
- Schindler, L. A., Burkholder, G. J., Morad, O. A., & Marsh, C. (2017). Computer-based technology and student engagement. *International Journal of Educational Technology in Higher Education*, 14(25).
- Schnotz, W., & Bannert, M. (2003). Construction and interference in multimedia learning. *Learning and Instruction*, 13(2), 141–156.
- Setyowati, B. L. (2020). Facebook's Covid-19 Pandemic Information Processing. *Journal of Communication and Media*, 01(01), 45–54.
- Skulmoski, G. J., Hartman, F. T., & Krahn, J. (2007). The Delphi method for graduate research. *International Journal of Project Management*, 25(1), 1–9. <https://doi.org/10.1016/j.ijproman.2006.10.007>
- Spante, M., Hashemi, S. S., Lundin, M., & Algers, A. (2018). Digital competence and digital literacy in higher education. *Computers in Human Behavior*, 81, 72–81. <https://doi.org/10.1016/j.chb.2017.12.006>
- Street, B. V. (2003). What's "new" in new literacy studies? *Current Issues in Comparative Education*, 5(2), 77–91.
- Sung, Y.-T., Chang, K.-E., & Liu, T.-C. (2016). The effects of integrating mobile devices with teaching and learning on students' learning performance: A meta-analysis. *Computers & Education*, 94, 252–275. <https://doi.org/10.1016/j.compedu.2015.11.008>
- Taber, K. S. (2018). The use of Cronbach's alpha. *Research in Science Education*, 48, 1273–1296.

- Unsworth, L. (2014). *Multimodal literacy and learning*. Cambridge University Press.
- van Deursen, A., Helsper, E., & Eynon, R. (2021). Development and validation of the Internet Skills Scale. *Information, Communication & Society*, 24(10), 1–17.
- van Laar, E., van Deursen, A., van Dijk, J., & de Haan, J. (2020). Determinants of 21st-century digital skills. *Computers in Human Behavior*, 72, 577–588.
- Walsh, M. (2010). Multimodal literacy: What does it mean? *Australian Journal of Language and Literacy*.
- Wineburg, S., & McGrew, S. (2017). Lateral reading and civic online reasoning. Stanford History Education Group.
- Zumbo, B. D. (2007). Three generations of validity. *Measurement: Interdisciplinary Research and Perspectives*.