A Multimodal HOTS-Oriented Worksheet Based On The PISA Reading Framework: Development and Validation Study in Middle School

Lia Sylvia Dewi, Vismaia S. Damaianti, Yeti Mulyati, And Rudi Adi Nugroho
Universitas Pendidikan Indonesia
Corresponding Author: liasylvdw@upi.edu

Abstract

This study developed and validated a digital multimodal student worksheet that integrates Higher-Order Thinking Skills (HOTS) and the PISA reading literacy framework for Indonesian middle school reading instruction. Although HOTS and literacy outcomes are emphasized in curriculum policy, teachers still have limited access to classroom-ready worksheets that *systematically operationalize* PISA cognitive processes (locating information, understanding, and evaluating/reflecting) within a single coherent learning tool. Addressing this gap, the study employed a Research and Development approach using the ADDIE model to design an LKPD that maps national competencies to PISA-oriented reading tasks supported by multimodal texts (e.g., infographics and charts). The product was reviewed by expert validators and piloted with seventh-grade students in a classroom setting. Validation results indicated that the LKPD met high standards of didactic, constructional, and technical quality and that the accompanying assessment items were appropriate for eliciting higher-order reading responses. The try-out data further showed that the LKPD can function diagnostically by revealing students' recurring challenges in inferential comprehension and information integration, while also supporting engagement with evaluative and reflective prompts through multimodal task design. Overall, the findings suggest that the developed LKPD is a valid and feasible tool for implementing HOTS-oriented, PISA-aligned reading activities and for informing instructional decisions based on students' domain-specific reading profiles. Future studies should examine broader implementation across diverse school contexts and evaluate learning impact over sustained use.

Keywords: digital worksheet, HOTS, multimodal literacy, PISA reading, reading literacy

Introduction

Reading literacy is a core skill essential for lifelong learning, critical participation in society, and success in academic and professional settings (Royce, 2021; Gorzycki et al., 2019; Toroujeni, 2022; Lehtonen, 2013). In the Indonesian education system, reading instruction is widely implemented across curricula; however, students continue to experience difficulties in interpreting and engaging with texts critically (Nandiasoka Annisawati & Ika Oktora, 2024; Damaianti et al., 2020; Tarisa Ariyani & Wahyuningsih, 2024). Classroom practices frequently emphasize recall and surface-level comprehension rather than encouraging deeper reasoning, textual analysis, or reflective response (Zunaedi et al., 2024; Segundo Marcos et al., 2020). Consequently, many students struggle to develop evaluative and inferential reading capacities required to navigate increasingly complex informational environments (Tarisa Ariyani & Wahyuningsih, 2024; Winahyu et al., 2020; Soisuwan et al., 2022).

The persistence of these challenges is reflected in international assessments such as the Programme for International Student Assessment (PISA). Indonesian students achieved an average reading score of 371 in PISA 2000, and this figure remained unchanged in the 2018 cycle, placing Indonesia 74th out of 79 participating countries. Although the 2022 cycle showed a slight increase to 379, Indonesian students still remained in the lowest performance quartile (OECD, 2023a). These results indicate ongoing difficulties in tasks that require understanding, evaluating, and reflecting on texts (Pratiwi, 2019). Prior studies suggest that such outcomes are closely related to limited exposure to higher-order reading tasks in everyday classroom instruction (Hidayati et al., 2020; Hu & Yu, 2021). Moreover, national initiatives such as the Asesmen Kompetensi Minimum (AKM) have not yet fully transformed instructional orientations toward higher-order literacy practices (Familiyana et al., 2022).

Ideally, reading education should cultivate students' ability to interact critically and reflectively with texts across genres and modalities (Medugno, 2020). This involves fostering reading experiences that promote reasoning, interpretation, and evaluation rather than passive consumption of information (Smith, 2012). In line with national curriculum directions, literacy instruction is expected to support the development of 21st-century skills, including critical thinking and creativity (Zaim & Zakiyah, 2024; BSKAP Kemendikbudristek, 2022). The integration of multimodal content—such as images, graphics, and digital media—is also increasingly emphasized to enhance engagement and contextual relevance. Studies have shown that multimodal texts can support comprehension by activating multiple cognitive channels and accommodating diverse learning preferences (Bowen, 2017; Danielsson & Selander, 2021).

Critical reading is closely linked to Higher-Order Thinking Skills (HOTS), which emphasize cognitive processes such as analyzing, evaluating, and creating as articulated in the revised Bloom's Taxonomy (Anderson et al., 2001). From this perspective, reading becomes an interactive process in which learners construct meaning, assess credibility, and evaluate arguments based on textual evidence (Borst & DiYanni, 2020). The PISA reading literacy framework further operationalizes these processes by organizing reading competence into locating information, understanding, and evaluating and reflecting, each corresponding to real-world literacy demands (OECD, 2023a). These domains provide a structured reference for designing instructional tasks that extend beyond literal comprehension.

Recent Indonesian studies consistently reveal persistent pedagogical challenges in implementing critical reading instruction. A systematic review by Setiawan et al. (2023) reported that Indonesian language teachers continue to face conceptual and instructional difficulties in fostering critical reading, particularly in structuring analytical and reflective processes. Similarly, textbook analyses indicate that learning activities remain largely focused on literal comprehension, with evaluative and inferential tasks unevenly distributed across instructional materials (Akbariski et al., 2024). Other studies further suggest that digital access alone does not automatically enhance critical reading competence without explicit instructional scaffolding (Yunaika, 2025; Fitrawati et al., 2023). Collectively, these findings indicate that challenges in critical reading instruction are rooted not only in curriculum design but also in the limited availability of pedagogically structured materials that support higher-order reading development.

Although previous studies have highlighted the potential of multimodal texts and HOTS-based approaches to enhance reading engagement and comprehension (Katoningsih & Sunaryo, 2020; Sari & Prasetyo, 2021; Kayati, 2022; Fan & Zhang, 2024; Hahnel et al., 2018), most existing instructional materials address these elements in isolation. Few studies have translated HOTS and PISA principles into concrete, classroom-ready worksheets that systematically integrate locating information, understanding, and evaluating or reflecting within a single coherent learning tool. As

a result, teachers have limited access to empirically validated instructional resources that operationalize higher-order reading processes in alignment with international literacy standards.

To address this gap, the present study develops and validates a digital multimodal student worksheet (LKPD) aligned with HOTS principles and the PISA reading literacy framework. Unlike conventional worksheets that emphasize basic comprehension or isolated higher-order questions, the developed LKPD integrates multimodal texts and HOTS-oriented tasks that explicitly map national learning competencies to PISA cognitive domains. By providing a validated instructional and diagnostic tool, this study seeks to support teachers in implementing higher-order reading instruction and in identifying students' reading literacy profiles within Indonesian middle school contexts.

Accordingly, this study aims to develop and validate a digital multimodal LKPD aligned with HOTS and the PISA reading literacy framework. Specifically, the study examines (1) the validity of the developed LKPD in terms of didactic, constructional, and technical aspects, and (2) its practicality as an instructional tool for supporting higher-order reading activities in Indonesian secondary education.

Literature Review

This literature review synthesizes key theoretical and empirical studies that inform the development of a multimodal student worksheet aligned with HOTS and the PISA reading literacy framework. The review focuses on four interrelated areas: critical reading as a core literacy competence, HOTS as a cognitive foundation for higher-level reading processes, the PISA reading literacy framework as a reference for authentic reading assessment, and the use of multimodal and digital worksheets in literacy instruction. These domains are examined because they directly shape the design principles, task structures, and assessment orientation of the developed LKPD. By integrating insights from these strands of literature, this review clarifies how critical reading competencies can be operationalized through HOTS-oriented and PISA-aligned tasks, and how multimodal worksheet design can support instructional and diagnostic purposes in classroom contexts. Collectively, this synthesis establishes the theoretical and empirical basis for the present study and positions it within broader efforts to develop context-sensitive and pedagogically grounded literacy materials.

Critical reading and higher-order thinking skills (HOTS)

Critical reading extends beyond decoding and literal comprehension, requiring readers to actively interpret meaning, evaluate arguments, and reflect on how texts position ideas and audiences (Graham-Matheson, 2018; Borst & DiYanni, 2020). In educational contexts, critical reading supports reasoned judgment and evidence-based interpretation rather than passive acceptance of textual information (Anuar & Sidhu, 2017; Arifin, 2020). This perspective implies that reading instruction should deliberately cultivate analytical, evaluative, and reflective practices, particularly when students encounter informational and digitally mediated texts that demand credibility appraisal (Hahnel et al., 2018).

Critical reading is conceptually aligned with Higher-Order Thinking Skills (HOTS), especially at the levels of analyzing, evaluating, and creating in the revised Bloom's Taxonomy (Anderson et al., 2001). Within this framework, reading activities should prompt learners to infer implicit meanings, justify interpretations, and synthesize information across textual elements.

However, HOTS in reading instruction should not be treated merely as an instructional label; rather, it must be operationalized through task designs that explicitly require reasoning processes and evidence use (Sani, 2019; Supriadi, 2020).

Empirical studies indicate that students frequently struggle with HOTS-related reading tasks, particularly those involving inferential comprehension and integration of information across paragraphs or representations (Hidayati et al., 2020; Aini, 2024). These difficulties suggest that challenges in reading performance are not solely due to limited language proficiency, but also to insufficient instructional opportunities that scaffold higher-order reasoning. In Indonesian instructional materials, HOTS-oriented tasks tend to emphasize analysis, while evaluative and creative processes are less consistently represented, resulting in limited practice with evidence-based judgment and synthesis (Damaianti et al., 2020; Ira, 2025).

A contributing factor to this imbalance is the continued reliance on traditional worksheets dominated by linear texts and closed-ended questions (Saputra et al., 2025; Setiawan et al., 2023). Such materials are effective for assessing recall and surface comprehension but provide minimal support for inferential reasoning and critical evaluation. Consequently, students often encounter HOTS terminology without sustained engagement in the cognitive processes that HOTS entails. This structural limitation highlights the need for instructional tools that embed higher-order reasoning demands within guided learning sequences rather than isolated test items.

The role of PISA in shaping literacy education

The Programme for International Student Assessment (PISA) plays a significant role in shaping contemporary perspectives on reading literacy by emphasizing the application of reading skills in real-world contexts. Unlike traditional school-based assessments that focus primarily on literal comprehension, PISA conceptualizes reading literacy as the capacity to understand, use, evaluate, and reflect on texts in order to achieve personal goals and participate effectively in society (OECD, 2019; OECD, 2023a). This orientation positions reading as an active, problem-solving process that requires higher-order cognitive engagement.

Within the PISA reading literacy framework, reading competence is organized into three core cognitive processes: locating information, understanding, and evaluating and reflecting (OECD, 2023a). The locating information process requires readers to search for and retrieve relevant information, often from non-linear or multimodal texts. The understanding process involves constructing meaning through integration of explicit and implicit information, while the evaluating and reflecting process represents the highest cognitive demand, requiring judgment, justification, and critical reasoning. These processes align closely with Higher-Order Thinking Skills (HOTS), particularly in terms of analysis and evaluation, making the PISA framework a relevant reference for HOTS-oriented reading instruction.

Empirical studies indicate that students experience different levels of difficulty across PISA reading domains. Difficulties in locating information frequently arise when texts are lengthy, multimodal, or contain competing sources of information, requiring strategic scanning and selective attention (Novela et al., 2022; Ramadhani & Zaim, 2023). Challenges in the understanding domain are commonly associated with inferential comprehension and the integration of ideas across paragraphs or representations, which remain persistent issues in Indonesian secondary education contexts (Aini, 2024; Setiawan et al., 2023). In the evaluating and reflecting domain, students are expected to demonstrate evidence-based judgment rather than personal opinion, yet research suggests that classroom reading tasks often provide limited opportunities to practice such reasoning systematically (Akbariski et al., 2024; Sirén & Sulkunen, 2025).

Despite the prominence of PISA in international literacy discourse, alignment between classroom instruction and PISA reading demands remains limited. Studies have shown that school-based reading activities tend to emphasize recall and surface comprehension, while higher-order processes such as evaluation and reflection are unevenly integrated into instructional materials (Kusmayati et al., 2021; Dewi et al., 2023). As a result, students may demonstrate familiarity with reading tasks in classroom settings but struggle when confronted with PISA-style assessments that require reasoning across texts, modes, and contexts.

From an instructional perspective, the PISA reading literacy framework offers a pedagogical reference for designing reading materials that integrate higher-order thinking, multimodal interpretation, and authentic assessment. By adopting PISA-aligned cognitive processes as design principles, instructional tools such as student worksheets can function not only as learning materials but also as diagnostic instruments for identifying students' reading strengths and weaknesses. In this sense, PISA serves not merely as an external benchmark but as a framework that can inform classroom-level instructional design, particularly in efforts to strengthen critical reading and HOTS-oriented literacy practices.

The PISA reading literacy framework

The PISA reading literacy framework conceptualizes reading as the capacity to understand, use, evaluate, and reflect on texts in order to achieve personal goals and participate effectively in society (OECD, 2023a). Unlike traditional school assessments, PISA emphasizes real-world reading demands and organizes reading competence into three core cognitive processes: locating information, understanding, and evaluating and reflecting. Each process requires increasingly complex reasoning and strategic engagement with texts.

Research indicates that students experience distinct challenges across these domains. Difficulties in locating information often arise when texts are lengthy, multimodal, or contain competing information sources, requiring selective attention and strategic scanning (Novela et al., 2022; Ramadhani & Zaim, 2023). Challenges in the understanding domain are commonly associated with inferential processing and integration of implicit information, particularly in texts combining verbal and visual elements (Setiawan et al., 2023; Aini, 2024). The evaluating and reflecting domain demands the highest level of cognitive engagement, as students must assess credibility, justify judgments, and connect textual content with external knowledge (Sirén & Sulkunen, 2025).

Studies examining Indonesian classroom practices suggest that alignment between reading instruction and PISA literacy demands remains limited. Classroom reading tasks frequently emphasize literal comprehension and isolated question types, providing insufficient exposure to the reasoning processes emphasized in PISA assessments (Dewi et al., 2023; Kusmayati et al., 2021). As a result, students may perform adequately on school-based assessments but struggle when confronted with complex, PISA-style reading tasks. These findings underscore the importance of using the PISA framework not merely as an assessment benchmark but as a pedagogical reference for instructional design.

Multimodal texts and digital literacy

Contemporary literacy practices increasingly involve multimodal texts that integrate written language with images, graphics, and digital representations (Adami & Kress, 2014; Jewitt et al., 2016). Multimodal literacy requires readers to interpret relationships across modes and to

coordinate visual and verbal information to construct meaning. When such relationships are not explicitly addressed in instruction, students may misinterpret data displays or fail to integrate information across representations (Danielsson & Selander, 2021).

Empirical evidence suggests that multimodal texts can enhance engagement and comprehension by activating multiple cognitive channels and accommodating diverse learning preferences (Bowen, 2017; Kayati, 2022). Digital learning platforms further enable the design of interactive reading tasks that simulate authentic literacy contexts, including PISA-style assessments (Rahmadani & Putri, 2021; Zahara et al., 2021). However, research also cautions that technological integration alone does not guarantee improved critical reading outcomes without explicit instructional scaffolding (Akbarov & Alimova, 2024).

Digital student worksheets (LKPD) function as instructional scaffolds that guide learners through structured reading activities, combining content presentation, guided questions, and assessment tasks (Widjajanti, 2008). When designed with HOTS-oriented prompts and aligned with PISA reading domains, digital LKPDs have the potential to support both instruction and diagnostic assessment. Nevertheless, prior studies indicate that many LKPDs prioritize visual appeal or content coverage while giving limited attention to systematic validation of didactic coherence, cognitive alignment, and classroom practicality (Setiawan et al., 2023).

Digital learning tools and worksheet innovation

The rapid expansion of digital learning environments has accelerated the development of interactive instructional materials that support student-centered and flexible learning (Avdeeva & Tarasova, 2023). In literacy instruction, digital tools enable teachers to design learning activities that integrate text, visuals, and interactive features, thereby supporting students' engagement with complex reading tasks. One widely adopted instructional format in Indonesian classrooms is the student worksheet (Lembar Kerja Peserta Didik—LKPD), which functions as a scaffold to guide learners through structured learning experiences.

Digital LKPDs differ from conventional paper-based worksheets in that they allow for interactivity, immediate feedback, and the integration of multimodal content (Adilla et al., 2019). Prior studies indicate that electronic worksheets can enhance student engagement and support higher-order thinking when tasks are designed with clear instructional purposes and cognitive alignment (Widjajanti, 2008). In reading instruction, digital LKPDs offer opportunities to embed HOTS-oriented prompts, varied response formats, and authentic reading tasks that resemble large-scale literacy assessments.

Technological platforms such as Liveworksheet and Canva provide practical affordances for developing visually rich and interactive LKPDs. These platforms enable the incorporation of multiple question types, including multiple-choice, matching, short-answer, and open-ended responses, which are essential for assessing diverse reading processes (Zahara et al., 2021; Rahmadani & Putri, 2021). Research has shown that students respond positively to digital worksheets that combine interactivity with multimodal representations, particularly when visuals are used to clarify abstract or complex textual information (Kayati, 2022).

However, several studies caution that innovation in digital format alone does not guarantee improved literacy outcomes. Without pedagogical structure and cognitive scaffolding, digital worksheets may replicate the limitations of traditional materials by emphasizing surface-level comprehension or task completion rather than higher-order reasoning (Setiawan et al., 2023). Furthermore, many LKPDs are developed primarily for content delivery or visual appeal, while their didactic coherence, technical functionality, and alignment with higher-order literacy

frameworks are rarely examined through systematic validation processes (Widjajanti, 2008; Adilla et al., 2019).

Recent research also highlights the importance of aligning digital worksheets with national and international literacy assessment frameworks. Integrating PISA-oriented reading tasks and HOTS indicators into LKPD design can help bridge the gap between classroom instruction and the cognitive demands of large-scale literacy assessments (Kusmayati et al., 2021; Dewi et al., 2023). In this context, digital LKPDs can serve not only as instructional materials but also as diagnostic tools that provide insight into students' reading literacy profiles.

Taken together, the literature suggests that effective digital worksheet innovation requires more than technological enhancement. It demands careful integration of cognitive objectives, multimodal design, and validated assessment principles. These considerations underscore the need for developing and validating digital multimodal LKPDs that operationalize HOTS and PISA reading domains within a coherent instructional framework, as pursued in the present study.

Theoretical perspectives on reading comprehension questions

Reading literacy frameworks emphasize that comprehension should be assessed through question types that reflect different levels of cognitive processing and meaning construction. In literacy education, reading comprehension questions function not only as assessment tools but also as instructional prompts that shape how students interact with texts (Canning, 2017). Therefore, the theoretical basis of question design is central to the development of valid and meaningful reading assessments.

Prior research has categorized reading comprehension questions into several types, including identifying main ideas, understanding details, interpreting vocabulary in context, sequencing events, making inferences, and evaluating or creating responses (Damaianti, 2021). These categories correspond to increasing levels of cognitive demand and provide indicators of students' depth of text processing. In particular, inference-based and evaluative questions are widely recognized as key measures of critical reading because they require integration of textual information, prior knowledge, and judgment.

Inference questions stimulate higher-level reasoning by prompting readers to derive implicit meanings that are not explicitly stated in the text. Such questions require students to connect ideas across sentences or paragraphs and to construct coherent interpretations based on textual cues. However, empirical studies indicate that inference-based comprehension remains a persistent difficulty for many learners, especially when tasks demand evidence integration rather than surface-level guessing (Hidayati et al., 2020; Aini, 2024). This challenge underscores the importance of designing inference questions that are supported by clear prompts and aligned with instructional scaffolding.

Evaluative and reflective questions represent the highest level of reading comprehension, as they require students to assess the credibility, relevance, or quality of information and to justify judgments using textual evidence (Damaianti et al., 2020). In international literacy frameworks such as PISA, evaluation and reflection are not treated as opinion-based responses but as evidence-based reasoning processes that involve weighing alternatives, monitoring credibility, and connecting textual ideas to external criteria (OECD, 2019; OECD, 2023a). Consequently, reading assessments must ensure that evaluative questions include explicit justification requirements rather than allowing purely subjective responses.

From a measurement perspective, reading comprehension questions can be implemented through varied formats, including multiple-choice, short constructed responses, and extended

open-ended responses. Each format serves a different diagnostic purpose. Multiple-choice items are effective for assessing locating information and basic understanding, while open-ended responses are more suitable for capturing inferential reasoning and evaluative judgment (Damaianti, 2021). Therefore, a combination of item types is necessary to represent the full range of cognitive processes involved in reading literacy.

In the context of worksheet-based instruction, comprehension questions should be embedded within learning sequences that guide students from surface comprehension toward deeper analysis and evaluation. When question types are not theoretically aligned with targeted cognitive processes, students may succeed in answering items without engaging in meaningful reasoning. This misalignment weakens both instructional and assessment validity. Accordingly, the development of HOTS-oriented and PISA-aligned LKPDs requires careful consideration of question typology, cognitive demand, and response format to ensure that higher-order reading processes are consistently elicited.

The present study draws on these theoretical perspectives to design reading comprehension questions that align with PISA reading domains and HOTS indicators. By integrating multiple question formats and explicitly targeting locating information, understanding, and evaluating/reflecting processes, the developed LKPD aims to function both as an instructional scaffold and as a diagnostic tool for assessing students' critical reading competencies.

Methodology

Research design and approach of the study

This study adopted a Research and Development (R&D) approach using the ADDIE model—Analysis, Design, Development, Implementation, and Evaluation—to guide the systematic development of a digital multimodal student worksheet (LKPD) (Sugiyono, 2023). The ADDIE model was selected because it supports iterative product design, validation, and revision in authentic instructional contexts, making it suitable for research focused on developing and validating instructional materials rather than examining causal effects (Mulyatiningsih, 2016; Baharuddin, 2018).

The primary objective of this study was to develop an LKPD aligned with Higher-Order Thinking Skills (HOTS) and the PISA reading literacy framework. Accordingly, the research emphasized product validity, practicality, and functional applicability. Data were collected and analyzed across the ADDIE stages using both quantitative and qualitative methods. Expert validation data were analyzed descriptively using percentage criteria and Aiken's V to examine didactic, constructional, and technical validity. Classroom try-out data were used to examine practicality in terms of clarity, usability, and task feasibility, drawing on student questionnaires, teacher feedback, and classroom observations.

Qualitative data obtained from interviews and observations were analyzed thematically to provide contextual insights into the implementation process. Open-ended responses were scored by two independent raters using a shared rubric based on HOTS and PISA descriptors, with prior calibration conducted to ensure scoring consistency. Data triangulation across expert review, classroom implementation, and student performance was employed to enhance the credibility of the findings. Within this R&D framework, effectiveness was interpreted functionally as the LKPD's capacity to elicit higher-order reading responses rather than as evidence of instructional impact.

Research site and participants

The study was conducted at SMP Tridaya Tunas Bangsa, a private junior secondary school located in Cimahi, West Java. The school was purposively selected because it represents students within the age range relevant to the PISA target population and has prior experience implementing digital learning activities. In addition, the school administration demonstrated institutional support for pedagogical innovation and facilitated access for classroom observations and product trials.

The participants consisted of 33 Grade VII students aged 13–14 years, including both male and female students, who were selected through purposive sampling. The selection criteria focused on students' basic reading proficiency and familiarity with digital learning platforms, as these competencies were required to engage with the multimodal LKPD. All participating students had prior exposure to digital learning environments such as Google Classroom and Liveworksheet through regular classroom instruction, ensuring the feasibility of the digital worksheet implementation.

In addition to students, one Indonesian language teacher and one school principal were involved in the study. The teacher acted as the classroom implementer during the try-out stage and contributed practical feedback based on instructional experience. The school principal functioned as an institutional gatekeeper and provided contextual information related to school policies and learning conditions during the needs analysis stage. For expert validation, the LKPD was reviewed by two university lecturers with expertise in literacy education and instructional design, and one practicing teacher. Furthermore, two literacy experts evaluated the open-ended test instruments to ensure content relevance and pedagogical alignment.

Data collection procedures

Data collection was carried out in three main stages. Ethical approval for the study was obtained from the school administration prior to data collection. Informed consent was secured from the Indonesian language teacher, the school principal, and students' parents or guardians. Participation was voluntary, and data confidentiality was ensured.

1. Needs Assessment: This stage was conducted through semi-structured interviews with the Indonesian language teacher and the school principal, as well as student surveys. The instruments focused on identifying instructional challenges, students' literacy needs, and current usage of LKPDs in reading instruction. The results informed the initial design specifications of the LKPD.

2. Product Design and Development: A digital, multimodal LKPD was developed using the Liveworksheet platform. The worksheet integrated PISA-style reading items, including multiple-choice, matching, short-answer, and open-ended questions. The content design emphasized Higher-Order Thinking Skills (HOTS) and incorporated multimodal texts such as infographics and charts to support reading tasks.

3. Validation Process:
    a) Product Validation: The LKPD was evaluated by three experts consisting of two university lecturers with expertise in literacy education and instructional design, and one practicing Indonesian language teacher with more than five years of teaching experience. A structured validation instrument was used to assess didactic quality, construction logic, and technical functionality using a four-point Likert scale.

b) Instrument Validation: The open-ended literacy test items were validated using a structured validation rubric completed by two literacy experts. The rubric assessed item relevance, clarity of wording, alignment with HOTS indicators, and correspondence with PISA reading domains. Each criterion was rated on a four-point scale, and the validation results were analyzed using Aiken's V formula to determine the content validity of each item.

Data analysis

Product validation by student

Product validity data obtained from expert judgments were analyzed using percentage scores. Each validator's score was divided by the maximum possible score and multiplied by 100 to obtain a validity percentage, which was then interpreted using predetermined validity categories adapted from (Riduwan, 2010) as shown in Table 1.

$$\text{Validation Score (\%)} = \left(\frac{\text{Total Score}}{\text{Maximum Score}} \right) \times 100 \quad (1)$$

Table 1. *Category interpretation*

Percentage (%)	Category
81–100	Very Valid
61–80	Valid
41–60	Fairly Valid
21–40	Less Valid
0–20	Not Valid

Notes: adapted from Riduwan (2010)

Instrument validation (Aiken's V)

The content validity of open-ended test items was calculated using Aiken's V to measure the level of agreement among expert validators. Each item was rated based on relevance and clarity, and the Aiken's V coefficient was computed using the standard formula and interpreted according to established validity criteria (Aiken, 1985), as presented in Table 2.

$$V = \frac{\Sigma s}{n(c-1)}$$

- $s = r - l0s = r - l\_0s = r - l0$ *(rating minus lowest score),*
- $rrr = score given by validator,$
- $l0l\_0l0 = lowest possible score,$
- $ccc = highest possible score,$
- $nnn = number of validators.$

.

Tabel 2. *Interpretation*

Aiken's V Value	Validity Level
≥ 0.80	High
0.40–0.79	Moderate
< 0.40	Low

Note. Adapted from (Aiken, 1985)

Item difficulty (objective tests)

The difficulty index of multiple-choice and matching items was calculated by dividing the number of correct responses by the total number of participants. The resulting index values were interpreted using difficulty level criteria adapted from (Mukrimaa et al., 2022), as shown in Table 3.

$$P = \frac{NP}{N}$$

- *PPP = item difficulty,*
- *BBB = number of correct responses,*
- *NNN = total number of participants.*

Tabel 3. *Difficulty levels*

Difficulty Index	Level
0.71–1.00	Easy
0.31–0.70	Medium
0.00–0.30	Hard

Notes. Adapted from (Mukrimaa et al., 2022)

Findings

Alignment with PISA literacy domains and national competencies

The development process aligned the PISA reading literacy domains—locating information, understanding, evaluating, and reflecting—with national curriculum learning outcomes (OECD, 2018). Table 4 shows the integration of literacy indicators with relevant text types and cognitive processes.

Validation of the LKPD and Assessment Instruments

Table 4. *Alignment of learning competencies with PISA reading domains*

Competency Code	PISA Literacy Domain	Indicators	Text Type
A.1.1	Locating Information	Ability to scan and identify relevant information within single or multiple texts.	Descriptive, Expository, Argumentative, Instructional, Transactional
A.1.2	Understanding; Evaluating & Reflecting	Comprehending, using, and evaluating fiction texts.	Narrative

| A.1.3 | Understanding | Comparing and integrating information across texts. | Descriptive, Narrative, Expository, Argumentative |
| A.1.4 | Evaluating and Reflecting | Assessing, predicting, and interpreting content within various genres. | Descriptive, Expository, Argumentative, Instructional |

This mapping shows that the developed LKPD fosters comprehensive literacy engagement with both informational and literary texts. The integration of diverse text types and skills aligns the product with PISA's emphasis on cognitive complexity and textual diversity.

Table 5. *Summary of LKPD product validation results*

Assessment Aspect	Validator 1	Validator 2	Validator 3	Mean (%)	Category
Didactic	100%	100%	93.3%	97.7%	Very Valid
Construction	100%	96.6%	90%	95.5%	Very Valid
Technical	100%	88%	92%	93.3%	Very Valid
Average	100%	94.8%	91.7%	95.5%	Very Valid

The overall validity score of 95.5% indicates that the multimodal LKPD met all three expert criteria at a "very valid" level. Reviewers also highlighted that the tasks and question formats effectively incorporated HOTS elements and aligned with PISA-style problem-solving structures.

Student performance analysis

Validation of the 35 multiple-choice and short-answer items was conducted using SPSS with the *Point Biserial Correlation (rPBIS)* method. The results indicated that most items achieved valid status, while a few required revisions due to weak discrimination values.

Table 6. *Summary of objective test validation results*

Category	Valid Items	Invalid Items
Locating Information	2, 3, 4, 6, 7, 8, 9, 10	1, 5
Understanding Text	12, 15–19, 22–23, 25–26, 28	11, 13, 14, 20, 21, 24, 27
Evaluating and Reflecting	29–33	34–35

Out of 35 items, 28 were valid and 7 required revisions. The valid items reflected strong alignment with PISA task distribution, emphasizing inferential and evaluative comprehension. Five open-ended reading questions were validated using Aiken's V, focusing on completeness, linguistic clarity, and cognitive depth. As shown in Table 4, all five items achieved a high validity index (V ≥ 0.80), confirming their appropriateness for assessing higher-order reading skills.

Table 7. *Summary of open-ended test validation results*

Item No.	Average Validity Coefficient	Interpretation
36	0.8125	High
37	0.8625	High
38	0.8625	High

39	0.875	High
40	0.8875	High

These results confirm that all constructed items are highly valid and aligned with the evaluating and reflecting domains of the PISA framework, integrating content, form, and linguistic precision.

Table 8. *Distribution of students' critical thinking levels*

Score Range	Category	Percentage of Students
>86	Excellent (Proficient)	0%
76–85	Good (Competent)	9.3%
65–75	Fair	33.3%
55–65	Poor (Needs Intervention)	15%
<54	Very Poor (Needs Intensive Support)	42%

This underperformance may be attributed to the lack of HOTS-oriented instruction in Indonesian classrooms, where students often engage with low-level recall tasks rather than interpretive or evaluative ones. (OECD, 2019) data reinforces this observation by showing that Indonesian students continue to perform below the OECD reading average. Analysis of item difficulty (Table 8) provides insight into students' strengths and weaknesses in processing text information.

Table 9. *Item difficulty distribution*

Range	Category	Total Items	Example Items (Domain)
0.00–0.32	Difficult	4	Items 5 (Locating Information), 11 & 14 (Understanding), 24 (Understanding)
0.33–0.66	Moderate	26	Items 1–4, 6–10 (Locating Information); 12–23 (Understanding); 30, 32–34 (Evaluating & Reflecting)
0.67–1.00	Easy	5	Items 7, 27, 29, 31, 35 (Evaluating & Reflecting)

Discussion

Interpretation of findings and student performance

The multimodal HOTS-oriented LKPD provided a diagnostic profile of students' critical reading across PISA domains rather than evidence of learning gains, which is consistent with its design purpose as an assessment-informed instructional tool. The distribution of performance—42% in the "very poor" category and only 9.3% in the "good" category—suggests that many students are not yet ready for sustained higher-order reading demands and may require systematic scaffolding. This pattern aligns with concerns raised in the PISA assessment framework and recent results emphasizing complex literacy demands such as evaluation, reflection, and evidence-based interpretation in authentic contexts (OECD, 2019, 2023a, 2023b). The domain-level variation also echoes recent Indonesian-focused syntheses showing persistent challenges in inferential comprehension and cross-text integration among secondary students (Aini, 2024; Setiawan et al., 2023). Therefore, the findings should be interpreted as mapping students' current cognitive readiness and opportunity-to-learn conditions rather than as proof that the LKPD directly improved performance.

The uneven performance across domains indicates that tasks requiring locating information and understanding become substantially more difficult when they involve comparison, inference, or synthesis rather than literal retrieval. Recent studies similarly report that learners struggle when comprehension requires constructing meaning across multiple segments or coordinating evidence, particularly in cognitively demanding reading situations (Aini, 2024; Akbarov & Alimova, 2024). In contrast, relatively stronger performance in evaluating and reflecting may reflect the accessibility of judgment-oriented prompts that allow intuitive or experience-based responses. Evidence from critical literacy research in PISA contexts suggests that students can appear successful on evaluation tasks when justification requirements are weak, even while deeper evidence integration remains limited (Sirén & Sulkunen, 2025). Consequently, the observed strengths in evaluation and reflection should be treated as emerging engagement rather than definitive mastery of critical literacy competence.

A key strength of the developed LKPD lies in its explicit alignment between curriculum objectives and PISA-oriented cognitive processes, which helps teachers make higher-order reading requirements visible at the task level. The multimodal task format is also pedagogically relevant because contemporary literacy increasingly involves interpreting and integrating information from multiple representations, and such integration is central to PISA-aligned reading demands (OECD, 2023a; Sirén & Sulkunen, 2025). Moreover, recent literature indicates that multimodal materials can support engagement and reduce perceived difficulty when they provide meaningful contexts and clear task purposes, which may partially explain why students responded more readily to reflective prompts (Fan & Zhang, 2024). The LKPD's domain-based profiling further supports diagnostic use by distinguishing performance that stems from literal retrieval from performance that requires inference and evidence-based reasoning. In this sense, the product contributes a practical bridge between assessment literacy and instructional design, echoing calls for structured resources that operationalize critical reading rather than merely advocating it (Ramadhani & Zaim, 2023; Setiawan et al., 2023).

Despite these strengths, the LKPD may fall short in consistently eliciting rigorous evidence-based reasoning, especially if some evaluative items permit subjective responses without requiring explicit textual or multimodal justification. This limitation is important because recent critical literacy research emphasizes that evaluation and reflection in PISA contexts should involve credibility monitoring, evidence tracing, and integration across information sources rather than opinion-only judgments (OECD, 2023a; Sirén & Sulkunen, 2025). In addition, if item scaffolds are uneven across domains, the observed domain differences may partly reflect task design features (e.g., prompt clarity, modality complexity, or justification cues) rather than students' underlying competence alone. The product is also constrained by the scope of its current implementation, since a single administration and a specific student population limit generalizability and do not capture growth over time. Therefore, transparency requires acknowledging that the LKPD functions well as a diagnostic snapshot but still needs refinement to strengthen construct representation of higher-order evidence use, especially within multimodal evaluation tasks (Ramadhani & Zaim, 2023; Setiawan et al., 2023)

For classroom practice, teachers can implement multimodal LKPDs most effectively by sequencing tasks from guided analysis toward independent evaluation, ensuring that students first learn how to locate, connect, and verify evidence before producing judgments. Practical scaffolds should include prompts such as "Which sentence/visual supports your claim?" "What evidence contradicts it?" and "How do two representations jointly explain the issue?" because these cues operationalize HOTS as observable reasoning actions rather than abstract labels. Schools can support implementation by strengthening digital literacy routines, including how students interpret

tables, charts, and online information credibility, which is increasingly emphasized in contemporary literacy assessment methodologies (Avdeeva & Tarasova, 2023; OECD, 2023a). This recommendation is consistent with recent findings that digital formats alone do not automatically improve integration skills, so instruction must explicitly teach strategies for navigating and evaluating information (Akbarov & Alimova, 2024). Accordingly, the LKPD should be used not only as an assessment artifact but as a guided learning sequence where teacher feedback targets evidence use, cross-mode integration, and reasoning quality.

At the policy level, curriculum developers and administrators can integrate multimodal assessment strategies by defining clear task specifications that require evidence-based reasoning across modes, aligned with PISA-oriented competencies and national learning outcomes (OECD, 2023a, 2023b). Professional development policies should prioritize teachers' assessment literacy, including how to design prompts that elicit justification and how to evaluate student reasoning beyond correctness, reflecting recommendations from recent Indonesian critical reading syntheses (Setiawan et al., 2023). Infrastructure support is also essential, because multimodal and digital literacy tasks depend on stable access to devices, connectivity, and classroom routines that make multimodal reading feasible at scale. Equity-oriented planning is particularly important given evidence that ICT and digital literacy factors can relate to reading literacy outcomes in Indonesia, implying that resource gaps may widen achievement gaps if implementation is uneven (Nandiasoka Annisawati & Ika Oktora, 2024). Therefore, policy action should combine standards for multimodal HOTS assessment with pragmatic support for infrastructure, teacher capacity, and equitable access.

Future research should test the LKPD across diverse populations, school contexts, and proficiency levels to evaluate robustness, fairness, and generalizability of the domain profile patterns. Studies should also extend multimodal tasks to additional text genres (e.g., argumentative, informational, digital sources with hyperlinks) to examine whether the same strengths and weaknesses persist across literacy contexts, as recommended by recent work emphasizing real-world reading demands (OECD, 2023a; Sirén & Sulkunen, 2025). Experimental or quasi-experimental designs can investigate whether repeated LKPD use with explicit scaffolding produces measurable growth in inferential comprehension, evidence integration, and evaluation quality over time. Longitudinal research would be especially valuable to examine whether students' apparent gains in evaluative engagement develop into evidence-based critical literacy, addressing the concern that subjective judgment can mask weak analytical rigor (Fan & Zhang, 2024; Sirén & Sulkunen, 2025). Finally, future validation work should refine item design by strengthening justification requirements and calibrating multimodal complexity so that higher performance more reliably reflects higher-order reasoning rather than familiarity-driven responding (Ramadhani & Zaim, 2023; Setiawan et al., 2023).

Evaluation and reflection domain: strengths and underlying factors

Contrary to the initial assumption that higher-order reading would be uniformly difficult, students in this study displayed a distinctive pattern in the evaluation and reflection domain, where 42.9% of the objective items were categorized as easy. This relative ease suggests that evaluative prompts may have been more accessible than tasks emphasizing locating information and constructing understanding, even though they are classified as higher-order processes. A comparable tendency is implied by work on critical literacy in PISA-related assessment, where evaluation and reflection items often involve judging the purpose, stance, or trustworthiness of texts in ways that can invite intuitive responses (Sirén & Sulkunen, 2025). Evidence from

Indonesian textbook analyses also indicates that evaluative components can appear as recognizable task forms (e.g., judging claims or perspectives), even when deeper evidence-based reasoning is not consistently demanded (Akbariski et al., 2024). Therefore, the present results are plausibly consistent with recent findings showing that "evaluation" performance may reflect perceived familiarity with evaluative question formats rather than uniformly strong critical literacy.

One plausible explanation concerns the task affordances of evaluation and reflection prompts, which may reduce perceived cognitive burden when they allow students to draw on familiar contexts and personal experience. Research on reading engagement and achievement emphasizes that students' willingness to engage with a text and respond meaningfully is shaped by task accessibility and motivational conditions, not only by cognitive skill demands (Allington & McGill-Franzen, 2021). In addition, recent work in technology-mediated learning suggests that positive affect and engagement can facilitate participation and confidence, although engagement alone does not guarantee rigorous reasoning (Fan & Zhang, 2024). When evaluative items resemble everyday judgment tasks (e.g., "Is this fair?" "Do you agree?"), students may respond fluently even if their responses are not grounded in careful textual evidence. Accordingly, the observed ease should be interpreted as an interaction between task design, perceived familiarity, and engagement rather than as definitive proof of advanced evaluative proficiency.

From a cognitive standpoint, schema activation can facilitate quick evaluative judgments because prior knowledge and everyday reasoning provide readily available interpretive frames. In PISA-aligned reading contexts, evaluation and reflection are not limited to stating opinions but also involve monitoring credibility and integrating evidence, which distinguishes surface evaluation from critical literacy competence (OECD, 2019; Sirén & Sulkunen, 2025). Recent analyses caution that learners may appear competent on evaluative prompts when justification requirements are weak, while still struggling with analytical rigor such as identifying assumptions, reconciling multiple sources, or tracking evidence across a text. This distinction is consistent with findings that classroom and textbook practices may present evaluation-like tasks but do not always scaffold evidence-based reasoning systematically (Akbariski et al., 2024). Therefore, the relatively high performance in this domain should be treated as an indicator of emergent evaluative engagement, not as conclusive evidence of fully developed critical reading.

The pattern also aligns with a broader instructional reality: in many settings, students encounter more routine prompts that invite judgment than tasks that demand systematic inference or cross-information integration. Indonesian curriculum materials and related studies indicate that higher-level items are often concentrated at the analytic level, while fewer tasks require explicit evaluation with evidence and fewer still require synthesis or creation, limiting sustained practice in rigorous higher-order reasoning (Damaianti et al., 2020). Moreover, studies comparing digital and traditional reading contexts suggest that format changes do not automatically strengthen integration skills, which are central to deep comprehension and argument evaluation (Akbarov & Alimova, 2024). When evaluation prompts are not tied to explicit evidence-tracking, students may succeed through intuitive or experiential reasoning and still struggle on items that require integrating claims and data across representations. Consequently, the domain-level "strength" observed here may reflect a gap between evaluative participation and evaluative rigor, which depends heavily on item design features and justification demands.

These findings have direct instructional implications for the multimodal LKPD developed in this study, particularly for differentiating between subjective evaluative responses and evidence-based evaluation. Recent literature on PISA-oriented reading emphasizes that evaluation and reflection should be operationalized through prompts that require students to cite textual or visual evidence, weigh alternatives, and justify conclusions, rather than merely expressing agreement or

personal preference (OECD, 2019; Sirén & Sulkunen, 2025). Therefore, LKPD tasks in this domain should incorporate explicit justification scaffolds (e.g., "Which sentence/visual supports your judgment?" "What counterevidence weakens the claim?") and multimodal integration requirements to prevent superficial performance. This approach is consistent with recent calls in Indonesian critical reading research to move beyond descriptive understanding toward structured, evidence-based reasoning practices that teachers can implement consistently (Setiawan et al., 2023). Finally, future iterations of the LKPD can strengthen the evaluation-reflection domain by increasing the proportion of items that demand cross-mode evidence integration and argument appraisal, ensuring that higher scores indicate deeper critical literacy rather than familiarity-driven responding.

Overall performance and implications for pedagogical practices

The overall distribution of student performance indicates substantial challenges in meeting higher-order reading demands, particularly given that 42% of students were classified in the "very poor" category. Rather than reflecting deficiencies in the LKPD design, this pattern highlights students' limited prior exposure to cognitively demanding reading tasks. The findings suggest that many students are unfamiliar with sustained analytical engagement, especially when tasks require inference, integration of information, or evaluation based on textual evidence. This gap underscores a broader instructional issue in which reading activities frequently emphasize surface comprehension rather than higher-order processing aligned with PISA and HOTS frameworks.

From a pedagogical perspective, the findings indicate that the integration of higher-order thinking skills into reading instruction needs to be more systematic and continuous. The LKPD was able to reveal variations in students' performance across cognitive domains, suggesting its potential role as a diagnostic and instructional support tool. However, the results also imply that students require consistent opportunities to practice higher-order reading processes, rather than encountering them only through isolated assessment tasks. Reading instruction, therefore, should move beyond recall-oriented questioning toward activities that encourage students to analyze textual structure, evaluate arguments, and reflect on meaning using explicit criteria.

The findings also reinforce the pedagogical value of multimodal materials in supporting student engagement. The interactive and visual elements of the LKPD appeared to sustain students' attention and facilitate initial interaction with texts. Nevertheless, the discrepancy between students' relatively stronger performance on evaluative tasks and their difficulties with comprehension and inference suggests that multimodality alone is insufficient to foster deep cognitive development. Without structured guidance, students may respond intuitively rather than analytically, particularly in tasks that appear open-ended or subjective.

These results imply that teachers play a critical role in mediating the use of multimodal LKPDs. Effective implementation requires instructional scaffolding that gradually increases cognitive demand and supports students in articulating reasoning, justifying interpretations, and using textual evidence. Pedagogical strategies such as guided questioning, collaborative discussion, and structured written responses may help students transition from intuitive reactions to more systematic critical reading practices. In this sense, the LKPD should be viewed not as a standalone instructional solution but as a tool that supports teachers in designing and facilitating higher-order reading activities within regular classroom instruction.

Limitations of the study

Despite its contributions, this study has several limitations that should be considered when interpreting the findings. First, the study was conducted in a single urban junior secondary school, which limits the transferability of the results to other educational contexts. Students in this setting may have had greater access to digital resources than those in rural or under-resourced schools, potentially influencing their familiarity with multimodal learning environments. Future studies should involve more diverse school settings to examine how multimodal HOTS-oriented LKPDs function across different infrastructural and socioeconomic conditions.

Second, the implementation of the LKPD was limited to a short trial period. As a result, the study did not examine how sustained or repeated use of multimodal LKPDs might shape students' engagement with higher-order reading tasks over time. Longer-term investigations are needed to explore how continuous exposure to HOTS-oriented and PISA-aligned reading activities may support the gradual development of critical reading processes.

Third, although the LKPD and its assessment instruments were validated by experts, classroom implementation relied on a limited instructional context. Variations in teacher facilitation and instructional pacing may have influenced how students engaged with the tasks. As this study focused on product development and functional applicability, it did not aim to isolate instructional effects. Future research may consider broader implementation contexts or comparative designs to further explore instructional dynamics associated with multimodal LKPD use. This study focused on informational reading texts aligned with PISA literacy domains and did not extend multimodal design to other text genres, such as literary texts or multimedia narratives. Consequently, the findings should be interpreted within the scope of the selected text types. Further research is recommended to explore the applicability of multimodal HOTS-oriented LKPDs across a wider range of genres and literacy contexts.

Conclusion and Recommendations/Implications

This study focused on the development and validation of a digital multimodal student worksheet (LKPD) aligned with Higher-Order Thinking Skills (HOTS) and the PISA reading literacy framework within the context of Indonesian secondary education. The findings indicate that the developed LKPD was successfully designed in accordance with national curriculum expectations and PISA reading domains, namely locating information, understanding, evaluating, and reflecting. The systematic ADDIE-based development process ensured that learning objectives, task structures, and multimodal elements were coherently aligned with targeted cognitive processes.

The validation results demonstrate that the LKPD and its accompanying assessment instruments meet high standards of didactic, constructional, and technical quality. Expert judgments confirmed the appropriateness of content organization, task clarity, and technical functionality, while item analysis and Aiken's V coefficients indicate that the assessment components are suitable for eliciting higher-order reading responses. These findings suggest that the LKPD is a valid instructional resource that can be used to support the design and implementation of HOTS-oriented reading tasks in classroom settings.

Insights from the product try-out further show that the LKPD is practical for classroom use and capable of revealing students' critical reading profiles across PISA reading domains. Student performance patterns indicate persistent challenges in inferential comprehension and text

integration, alongside relatively stronger engagement with evaluative and reflective tasks, particularly when supported by multimodal texts. Rather than indicating direct instructional impact, these findings highlight the role of the LKPD as an instructional and diagnostic tool that can assist teachers in identifying students' initial reading competencies and planning appropriate learning strategies.

Taken together, the results of this study demonstrate that HOTS and PISA-oriented literacy principles can be meaningfully operationalized into validated classroom materials through a structured development process. The multimodal LKPD developed in this study offers a practical learning resource that supports critical reading instruction, assessment, and instructional decision-making. In terms of implications, teachers are encouraged to integrate multimodal and HOTS-based worksheets to facilitate higher-order reading engagement and to inform differentiated instructional planning. At the curricular level, the alignment of classroom materials with national literacy initiatives and international frameworks such as PISA may contribute to more coherent literacy practices. Future research may explore broader implementation contexts, longitudinal use, and the role of similar instructional tools in supporting instructional planning and assessment across diverse educational settings.

Disclosure statement

Acknowledgments

References

Adami, E., & Kress, G. (2014). Introduction: Multimodality, meaning making, and the issue of "text." In *Text and Talk* (Vol. 34, Number 3). https://doi.org/10.1515/text-2014-0007

Adilla, T. N., Silitonga, F. S., & Ramdhani, E. P. (2019). Pengembangan Electronic Lembar Kerja Peserta Didik (E-LKPD) Berbasis Guided Inquiry Materi Kelarutan dan Hasil Kali Kelarutan *(Development of guided inquiry-based electronic student worksheets on solubility and solubility product)*. *Jurnal FMIPA Universitas Maritim Raja Ali Haji.*, *10*(1), 38–51. https://doi.org/10.21608/pshj.2022.250026

Agustin Vera Dewi, Apriliya, S., & Nur, L. (2025). A Qualitative Needs Analysis of SQ3R-Based Digital Learning Worksheets in Primary School Reading Activities. *Jurnal Cakrawala Pendas*, *11*(3). https://doi.org/10.31949/jcp.v11i3.13448

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, *45*(1). https://doi.org/10.1177/0013164485451012

Aini, R. (2024). Students' Abilities and Difficulties in Comprehending English Reading Texts at Second Grade Students' of SMA Negeri 1 Sakra Timur. *Educatoria : Jurnal Ilmiah Ilmu Pendidikan*, *4*(2). https://doi.org/10.36312/educatoria.v4i2.260

Akbariski, H. S., Harsiati, T., & Basuki, I. A. (2024). Assessment of Critical Reading Ability Using Indonesian Language Textbooks for the 2013 Curriculum and the Merdeka Curriculum. *Jurnal Ilmu Pendidikan*, *30*(1). https://doi.org/10.17977/um048v30i1p19-28

Akbarov, J., & Alimova, F. (2024). Digital versus Traditional Reading Materials in English Language Learning: A Comparative Analysis of Student Perceptions and Outcomes. *Research Studies in English Language Teaching and Learning*, *2*(2). https://doi.org/10.62583/rseltl.v2i2.40

Allington, R. L., & McGill-Franzen, A. M. (2021). Reading Volume and Reading Achievement: A Review of Recent Research. *Reading Research Quarterly*, *56*(S1). https://doi.org/10.1002/rrq.404

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C. (2001). *A Taxonomy for Learning,. Teaching, and Assesing; A revision of Bloom's Taxonomy of Education* (A Bridge E). Longman.

Anuar, N., & Sidhu, G. K. (2017). Critical reading skills: A survey of postgraduate students' perspective of critical reading. *Pertanika Journal of Social Sciences and Humanities*, *25*(February).

Arifin, S. (2020). The Role of Critical Reading to Promote Students' Critical Thinking and Reading Comprehension. *Jurnal Pendidikan Dan Pengajaran*, *53*(3), 318–326. https://doi.org/10.23887/jpp.v53i3.29210

Avdeeva, S., & Tarasova, K. (2023). Digital Literacy Assessment: Methodology, Conceptual Model and Measurement Tool. *Voprosy Obrazovaniya / Educational Studies Moscow*, *2023*(2). https://doi.org/10.17323/1814-9545-2023-2-8-32

Baharuddin, B. (2018). ADDIE Model Application Promoting Interactive Multimedia. *IOP Conference Series: Materials Science and Engineering*, *306*(1). https://doi.org/10.1088/1757-899X/306/1/012020

Bezemer, J. (2021). Multimodality: A Guide for Linguists. *Research Methods in Linguistics*, 1–18. https://doi.org/10.5040/9781350043466.ch-012

Borst, A., & DiYanni, R. (2020). Critical Reading across the Curriculum: Volume 2: Social and Natural Sciences. In *Critical Reading across the Curriculum: Social and Natural Sciences*. https://doi.org/10.1002/9781119155317

Bowen, T. (2017). Assessing visual literacy: a case study of developing a rubric for identifying and applying criteria to undergraduate student learning. *Teaching in Higher Education*, *22*(6), 705–719. https://doi.org/10.1080/13562517.2017.1289507

BSKAP Kemendikbudristek. (2022). Tahapan Implementasi Kurikulum Merdeka di Satuan Pendidikan. In *Kemendibudristek*.

Canning, P. (2017). Text World Theory and real world readers: From literature to life in a Belfast prison. In *Language and Literature* (Vol. 26, Number 2). https://doi.org/10.1177/0963947017704731

Damaianti, V. S. (2021). *Literasi Membaca* (Reading literacy) (1st ed.). PT. Refika Aditama.

Damaianti, V. S., Abidin, Y., & Rahma, R. (2020). Higher order thinking skills-based reading literacy assessment instrument: An Indonesian context. *Indonesian Journal of Applied Linguistics*, *10*(2). https://doi.org/10.17509/ijal.v10i2.28600

Danielsson, K., & Selander, S. (2021a). Multimodal Texts in Disciplinary Education. In *Multimodal Texts in Disciplinary Education*. Springer International Publishing. https://doi.org/10.1007/978-3-030-63960-0

Danielsson, K., & Selander, S. (2021b). *Multimodal Texts in Disciplinary Education : A Comprehensive Framework.* https://library.oapen.org/bitstream/id/634abdda-f258-4fb5-9e79-083ad9c75bd1/978-3-030-63960-0.pdf

Dewi, L. S., Damaianti, V. S., & [Nama Penulis Lain]. (2023). *Identifikasi tingkat kesukaran instrumen tes membaca HOTS pada lembar kerja multimodal: Meninjau perspektif asesmen berorientasi PISA bagi siswa SMP (Identification of item difficulty of HOTS reading test instruments in multimodal worksheets: Reviewing PISA-oriented assessment perspectives for middle school students). Proceedings/Unpublished Conference Paper*

Familiyana, L., Hardjono, H. S., & Suryani, I. (2022). Persepsi Guru terhadap Soal Asesmen Kompetensi Minimum (AKM) Literasi Membaca di SMP *(Teachers' perceptions of minimum competency assessment (AKM) reading literacy items in Junior High Schools). Silampari Bisa: Jurnal Penelitian Pendidikan Bahasa Indonesia, Daerah, Dan Asing, 5*(1). https://doi.org/10.31540/silamparibisa.v5i1.1697

Fan, J., & Zhang, Q. (2024). From literacy to learning: The sequential mediation of attitudes and enjoyment in AI-assisted EFL education. *Heliyon, 10*(17), e37158. https://doi.org/10.1016/J.HELIYON.2024.E37158

Fitrawati, Syarif, H., Zaim, M., & Perrodin, D. D. (2023). The perceptions of tertiary students and lecturers regarding CLIL-based critical reading material employing interactive multimedia. *Indonesian Journal of Applied Linguistics, 12*(3), 598–611. https://doi.org/10.17509/ijal.v12i3.36838

Gorzycki, M., Desa, G., Howard, P. J., & Allen, D. D. (2019). "Reading Is Important," but "I Don't Read": Undergraduates' Experiences With Academic Reading. *Journal of Adolescent and Adult Literacy.* https://doi.org/10.1002/jaal.1020

Graham-Matheson, L. (2018). Critical reading and writing. In *Essential Theory for Primary Teachers.* https://doi.org/10.4324/9781315754604-12

Hahnel, C., Goldhammer, F., Kröhne, U., & Naumann, J. (2018). The role of reading skills in the evaluation of online information gathered from search engine environments. *Computers in Human Behavior, 78*, 223–234. https://doi.org/10.1016/j.chb.2017.10.004

Hasanah, U., Edwita, & Ahmad Januar. (2021). Pendampingan Guru Mengembangkan Assesment Kompetensi Minimum (Akm) Berorientasi Pisa Untuk Meningkatkan Kualitas Hasil Pembelajaran Di Sekolah Dasar Wilayah Kabupaten Bogor (Teacher Support in Developing PISA-Oriented Minimum Competency Assessments (AKM) to Improve Learning Outcomes in Elementary Schools in Bogor Regency) *Jurnal Abadimas Adi Buana, 5*(01). https://doi.org/10.36456/abadimas.v5.i01.a3634

Hidayati, M., Inderawati, R., & Loeneto, B. (2020). The Correlations among Critical Thinking Skills, Critical Reading Skills and Reading Comprehension. *English Review: Journal of English Education, 9*(1), 69–80.

Hu, J., & Yu, R. (2021). The effects of ICT-based social media on adolescents' digital reading performance: A longitudinal study of PISA 2009, PISA 2012, PISA 2015 and PISA 2018. *Computers and Education, 175.* https://doi.org/10.1016/j.compedu.2021.104342

Ira. (2025). cS BAHASA INDONESIA SMP KELAS IX BERDASARKAN *HIGHER ORDER THINGKING SKILLS* (HOTS). *SASTRANESIA: Jurnal Program Studi Pendidikan Bahasa Dan Sastra Indonesia, 13*(2). https://doi.org/10.32682/sastranesia.v13e3/26/

Jewitt, Carey, Bezemer, Jeff, & O'Halloran. (2016). Introducing Multimodality. In *Multimodal Communication* (Vol. 5, Number 2). https://doi.org/10.1515/mc-2016-0026

Katoningsih, S., & Sunaryo, I. (2020). PROGRAMME FOR INTERNATIONAL STUDENT ASSESMENT (PISA) AS READING LITERACY STANDARD: CRITICAL THINKING

SKILL IS PRIORITY. *Education, Sustainability And Society*, *3*(1). https://doi.org/10.26480/ess.01.2020.08.10

Kayati, A. N. (2022). Pemanfaatan teks multimodal dalam pembelajaran bahasa Indonesia untuk penguatan literasi peserta didik *(Utilizing multimodal texts in Indonesian language learning to strengthen students' literacy). SANDIBASA I (Seminar Nasional Pendidikan Bahasa Dan Sastra Indonesia I)*, *4*(April), 385–398.

Kusmayati, N. B., Mulyati, Y., & Musyarofah, L. (2021). Pelatihan pembuatan alat evaluasi berbasis high order thinking skills (HOTS) dan kompetensi PISA membaca bagi guru-guru SMK (Training on developing HOTS-based and PISA-oriented reading assessment tools for vocational school teachers). *DIMASATRA: Jurnal Pengabdian Kepada Masyarakat*, *1*(2), 79–88.

Lehtonen, M. (2013). Reading, literacy, and education. *CLCWeb - Comparative Literature and Culture*, *15*(3). https://doi.org/10.7771/1481-4374.2240

Medugno, M. (2020). Postcolonial Poetics: 21st-Century Critical Readings. *Interventions*, *22*(2), 296–299. https://doi.org/10.1080/1369801x.2019.1660204

Mukrimaa, S. S., Nurdyansyah, Fahyuni, E. F., YULIA CITRA, A., Schulz, N. D., د, غسان., Taniredja, T., Faridli, E. Miftah., & Harmianto, S. (2022). Penilaian Hasil Belajar Bahasa (Language learning outcome assessment). In *Jurnal Penelitian Pendidikan Guru Sekolah Dasar* (Vol. 6, Number August).

Mulyatiningsih, E. (2016). *Metode Penelitian Terapan Bidang Pendidikan*. 35,110,114,120,121.

Nandiasoka Annisawati, P., & Ika Oktora, S. (2024). How does ICT literacy influence reading literacy score in Indonesia: first attempt using spatial analysis approach. *Journal of Applied Research in Higher Education*, *16*(1). https://doi.org/10.1108/JARHE-10-2022-0322

Nora ulfiah, Desy Hanisa Putri, & Iwan Setiawan. (2025). Development of Electronic Worksheet Based on Project Based Learning using Liveworksheet to Improve Critical Thinking Skills. *Jurnal Penelitian Pembelajaran Fisika*, *16*(1). https://doi.org/10.26877/jp2f.v16i1.1328

Novela, G. T., Asrowi, A., & Widyastono, H. (2022). Student's Reading Literacy: Opportunities and Characteristic for Instructional Media Development. *Journal of Education Technology*, *6*(1). https://doi.org/10.23887/jet.v6i1.42843

OECD. (2018). PISA 2018 : Reading Literacy Framework. In *PISA 2018 Assessment and Analytical Framework*. OECD Publishing.

OECD. (2019). PISA 2018 Assessment and Analytical Framework. In *OECD Publishing* (OECD Punli).

OECD. (2023a). PISA 2022 assessment and analytical framework. In *OECD Publishing*.

OECD. (2023b). *PISA 2022 Results (Volume I)*. OECD Publishing. https://doi.org/10.1787/53f23881-en

Pratiwi, I. (2019). PISA Effect On Curriculum In Indonesia. *Jurnal Pendidikan Dan Kebudayaan*, *4*(1), 51–71. https://doi.org/10.24832/jpnk.V4i1.1157

Rahmadani, E., & Putri, F. A. (2021). Pengembangan Lembar Kerja Siswa Interaktif Menulis Deskriptive Text Menggunakan Media Wizer dalam Mempromosikan Higher Order Thinking Skills (HOTS) untuk Siswa Kelas 8 Di MTSN 2 Medan (Development of Interactive Student Worksheets for Writing Descriptive Texts Using Wizer Media to Promote Higher-Order Thinking Skills (HOTS) among Grade 8 Students at MTsN 2 Medan). *BAHAS*, *32*(1). https://doi.org/10.24114/bhs.v32i1.25442

Ramadhani, F., & Zaim, M. (2023). Characteristics of PISA Reading Literacy Assessment: How can it be Implemented at Senior High School? *AIP Conference Proceedings*, *2805*(1). https://doi.org/10.1063/5.0148381

Riduwan. (2010). Skala Pengukuran Variabel-variabel Penelitian (Measurement scales for research variables). In *Alfabeta*. Alfabeta.

Royce, J. (2021). From Literacy to Information Literacy: Reading for Understanding in the Real World. *IASL Annual Conference Proceedings*. https://doi.org/10.29173/iasl8057

Sani, R. A. (2019). Pembelajaran Berbasis HOTS (Higher OrderThinking Skills). In *Pembelajaran Berbasis HOTS (Higher OrderThinking Skills)*.

Saputra, W., Cahyani, I., & Sastromiharjo, A. (2025). Trends and Innovation in Digital Reading Assessment: A Decade of Bibliometric Insights. *Journal of Languages and Language Teaching*, *13*(1). https://doi.org/10.33394/jollt.v13i1.12182

Sari, D. M. M., & Prasetyo, Y. (2021). Project-based-Learning on Critical Reading Course to Enhance Critical Thinking Skills. *Studies in English Language and Education*, *8*(2). https://doi.org/10.24815/siele.v8i2.18407

Segundo Marcos, R. I., López Ferández, V., Daza González, M. T., & Phillips-Silver, J. (2020). Promoting children's creative thinking through reading and writing in a cooperative learning classroom. *Thinking Skills and Creativity*, *36*. https://doi.org/10.1016/j.tsc.2020.100663

Setiawan, A., Hang, N. T. T., Fauzan, F., & Derana, G. T. (2023). Critical reading research and its implications for critical reading skills for Indonesian language teachers: A systematic literature review. *BAHASTRA*, *43*(2). https://doi.org/10.26555/bs.v43i2.500

Sirén, M., & Sulkunen, S. (2025). Critical literacy in the PISA 2018 reading literacy assessment. *Scandinavian Journal of Educational Research*, *69*(2). https://doi.org/10.1080/00313831.2023.2287458

Smith, F. (2012). Understanding reading: A psycholinguistic analysis of reading and learning to read, sixth edition. In *Understanding Reading: A Psycholinguistic Analysis of Reading and Learning to Read, Sixth Edition*. https://doi.org/10.4324/9780203142165

Soisuwan, T., Sinturat, T., & Lekdamrongsak, P. (2022). The Investigation in Reading Problems and Reading Strategies of Undergraduate Students. *Journal of Social Science Studies*, *9*(2), 90. https://doi.org/10.5296/jsss.v9i2.20487

Sugiyono. (2023). Metode Penelitian Pendidikan (Kuantitatif, Kulaitatif, dan R&D, dan Penelitian Pendidikan). In *Alfabeta*.

Supriadi, G. (2020). *Penilaian Berbasis Higher Order Thinking Skill (Hots)*.

Tarisa Ariyani, P., & Wahyuningsih, S. (2024). *Reading Literacy In a Multiligual Context: Challenges and Strategies for Senior High School Students In Indonesia*. https://doi.org/10.24090/celti.2024.1034

Toroujeni, S. M. H. (2022). Computerized testing in reading comprehension skill: investigating score interchangeability, item review, age and gender stereotypes, ICT literacy and computer attitudes. *Education and Information Technologies*, *27*(2). https://doi.org/10.1007/s10639-021-10584-2

Widjajanti, E. (2008). Kualitas Lembar Kerja Siswa *(Quality of student worksheets). Makalah Seminar Pelatihan Penyususnan LKS Untuk Furu SMA/SMK Pada Kegiatan Pengabdian Kepada Masyarakat Jurusan Pendidikan FMIPA UNY*, *76*(3), 61–64.

Wiley, J., & Sons, I. (2017). Critical Reading Across the Curriculum. In *Critical Reading Across the Curriculum*. https://doi.org/10.1002/9781119154907

Winahyu, S. K., Boeriswati, E., & Murtadho, F. (2020). Reading the argument as assessment process Indonesian opinion text. *International Journal of Advanced Science and Technology*, *29*(7 Special Issue).

Yunaika, W. (2025). Peran Literasi Digital terhadap Kemampuan Membaca Kritis dalam Pembelajaran Bahasa Indonesia (Kajian Literature Review). *Hikamatzu : Journal of Multidisciplinary*, *2*(2).

Zahara, M., Abdurrahman, A., Herlina, K., Widyanti, R., & Agustiana, L. (2021). Teachers' Perceptions of 3D Technology-Integrated Student Worksheet on Magnetic Field Material: A Preliminary Research on Augmented Reality in STEM Learning. *IOP Conference Series: Earth and Environmental Science*, *1796*(1). https://doi.org/10.1088/1742-6596/1796/1/012083

Zaim, M., & Zakiyah, M. (2024). Can the Merdeka Belajar curriculum really improve students' reading literacy? *JOALL (Journal of Applied Linguistics and Literature)*, *9*(1). https://doi.org/10.33369/joall.v9i1.32173

Zunaedi, B. N. H., Frida Unsiah, & Khairil Azwar Razali. (2024). Implementing Content and Language Integrated Learning (CLIL) Through Group Discussion in Reading Comprehension Class. *Journal of Language and Literature Studies*, *4*(1), 221–231. https://doi.org/10.36312/jolls.v4i1.1826