

Assessing materials development: An expert evaluation of instructional materials developed by pre-service English teachers

Sukarno^{a,1*}, Sari Hidayati^{a,2}, Ihtiara Fitrianiingsih^{a,3}, Dewi Nur Widiyati^{a,4}, Shabrina Yumna Azhra^{a,5}, Megan Jinabe^{a,6}, Fitriya Dessi Wulandari^{a,7}, Herwin Dyastuti^{a,8}, Kauna Karima^{a,9}, Adivta Yudha Tama^{a,10}, Nita Maghfiratul Jannah^{a,11}, Fadhl Mohammed Awadh Gharamah^{b,12}, Yasir A. Alsamiri^{c,13}



^a Universitas Negeri Yogyakarta, Jl. Colombo No.1, Karang Malang, Caturtunggal, Yogyakarta, 55281, Indonesia

^b Yafea University College, University of Lahej, Yafea, Lahj Governorate, Republic of Yemen

^c Islamic University of Madinah, Abo Bakr Al Siddiq, Al Jamiah, Madinah 42351, Saudi Arabia

¹ sukarno@uny.ac.id, ² sari_hidayati@uny.ac.id, ³ ihtiara.f@uny.ac.id, ⁴ dewinurwidiyati@uny.ac.id, ⁵ shabrinayumnaazhra@uny.ac.id,

⁶ meganjinabe.2023@student.uny.ac.id, ⁷ fitriyadessi.2023@student.uny.ac.id, ⁸ herwindyastuti.2023@student.uny.ac.id,

⁹ kaunakarima.2023@student.uny.ac.id, ¹⁰ adivtayudha.2023@student.uny.ac.id, ¹¹ nitamagfiratuljannah@uny.ac.id,

¹² fadhlgharamah2021@gmail.com, ¹³ yalsamiri@iu.edu.sa

*Corresponding author

ARTICLE INFO

ABSTRACT

Article history

Received 19 November 2025

Revised 18 March 2026

Accepted 02 April 2026

Available Online 30 April 2026

Keywords

Higher-order thinking skills

Instructional design proficiency

Materials development

Pedagogical principles

Pre-service English teachers

The ability of pre-service English teachers to create high-quality, contextually appropriate instructional materials is a critical component of effective pedagogy. This study provides a quantitative and analytical assessment of their proficiency in instructional design. By evaluating 28 material packages developed by pre-service teachers, this study identifies their strengths and remaining challenges in implementing modern pedagogical principles. A panel of ten subject experts evaluated the materials using a validated rubric assessing curriculum alignment, relevance, higher-order thinking skills (HOTS), and 21st-century skills. The results revealed variation in design competencies, with pre-service teachers performing best in foundational areas such as curriculum alignment (mean score 7.875) and pedagogy (mean scores for Suitability and 21st-century skills were 7.679 and 7.357, respectively). However, the study also found challenges in applying advanced pedagogical concepts, such as HOTS, Industry 4.0, and Society 5.0, with mean scores of 7.000, 7.000, and 6.964, respectively. These findings suggest an area for further strengthening in teacher education programs, particularly in supporting the integration of complex, abstract concepts like HOTS and contemporary pedagogical frameworks. Without sufficient preparation in these areas, pre-service teachers may find it challenging to apply these concepts effectively in their instructional practice.



© The Authors 2026. Published by Universitas Ahmad Dahlan.

This is an open access article under the [CC-BY-SA](#) license.



How to Cite: Sukarno, Hidayati, S., Fitrianiingsih, I., Widiyati, D. N., Azhra, S.Y., Jinabe, M., Wulandari, F.D., Dyastuti, H., Karima, K., Tama., A.Y., Jannah, N.M., Gharamah, F.M.A., & Alsamiri, Y.A. (2026). Assessing materials development: An expert evaluation of instructional materials developed by pre-service English teachers. *English Language Teaching Educational Journal*, 9(1), 44-56. <https://doi.org/10.12928/eltej.v9i1.14997>

1. Introduction

Every educational system worldwide depends fundamentally on the quality and competence of its teachers. In the context of English language education, this role is particularly critical due to English's status as a global medium of communication (Julita, 2024; Permana et al., 2023). In Indonesia, the success of English as a Foreign Language (EFL) education is closely linked to teachers' ability to design and develop high-quality instructional materials that are responsive to learners' needs (Febriyanti, 2018). Recent practice-oriented studies have consistently shown that teachers' design decisions significantly influence student engagement and learning outcomes (Setyaningrum & Putri, 2023; Waluyo et al., 2023).

Instructional materials function as a crucial bridge between curriculum objectives and learners' needs. Well-designed materials are associated with improvements in engagement, comprehension, and cognitive development, including vocabulary acquisition, reading comprehension, and motivation (Azhra et al., 2024; Zein et al., 2020). This is particularly evident in studies examining gamified and digital materials in Indonesian and regional contexts (Amalia et al., 2023; Maming et al., 2023; Sumarni et al., 2023). Consequently, the ability to develop effective instructional materials constitutes a core competence for pre-service English teachers, as reflected in studies exploring their experiences in material design and the integration of digital tools such as Quizizz (Mat et al., 2025; Ölçer-Çevik & Kozaer, 2025).

From a theoretical perspective, constructivist learning theory and competency-based education have strongly informed recent approaches to instructional material design (Obasi, 2022). These perspectives emphasize active learning, authentic tasks, and the development of 21st-century competencies, including higher-order thinking skills (HOTS) and the 4Cs (Stofflett & Stoddart, 1994). Empirical studies in Indonesia and similar contexts have aligned instructional design practices with task-based and gamified approaches to support authentic learning experiences (Liu et al., 2025; Pratiwi & Waluyo, 2022). At the same time, research on digital tools such as Quizizz and Kahoot highlights their potential to enhance engagement and learning outcomes (Hamzah, 2023; Husnah et al., 2023; Maming et al., 2023).

However, despite these advances, a critical limitation persists in the literature. Existing studies focus predominantly on learner outcomes, perceptions, or the effectiveness of digital tools, rather than systematically examining the instructional materials as concrete artifacts of teacher competence (Bulqiyah & Munir, 2023). As a result, the pedagogical quality and coherence of teacher-developed materials remain underexplored.

This limitation is further reflected in recent practice-oriented studies that report positive effects of digital tools on engagement and vocabulary acquisition while also identifying challenges such as technological constraints and the difficulty of translating theory into effective instructional design (Manipatruni et al., 2023; Nugraheni & Fakhurriana, 2023). Taken together, these findings point to a significant gap: although digital tools are widely recognized for their pedagogical potential, there is still insufficient systematic, artifact-based evaluation of the instructional materials that teachers design and implement in EFL classrooms (Setyaningrum & Putri, 2023).

In the Indonesian context, this gap is particularly important given the strong emphasis on teacher competence as a determinant of educational quality. Contemporary studies consistently highlight the need for teachers to develop materials that integrate 21st-century skills, higher-order thinking, and culturally relevant content (Handayani et al., 2024; Hoir et al., 2019; Kapoyannis, 2021). Nevertheless, much of the existing research continues to rely on self-reported data or general descriptions of teaching practices, limiting a deeper understanding of pre-service teachers' actual capabilities (Haerazi, 2023; Thohir et al., 2021). This reliance hampers the ability to assess professional readiness in a concrete and evidence-based manner (Kao, 2023; Ridwan & Athena, 2023; Tortorelli et al., 2021).

Addressing this gap, the present study adopts a structured and systematic approach to evaluating instructional materials developed by pre-service English teachers. While prior research has examined instructional artifacts, relatively few studies have applied a comprehensive, multi-dimensional evaluation framework to assess their pedagogical coherence and alignment with professional competencies. This study responds to that need by employing expert-based analysis using the SQRAR model. The framework enables a systematic examination of key dimensions of instructional quality, including structural organization, pedagogical alignment, task design, and assessment coherence. By

operationalizing these elements, the study moves beyond surface-level evaluation to capture the depth and integration of pedagogical decision-making reflected in instructional materials (Esteban & Sung, 2025; Gorozhankina & Grechukhina, 2022). This aligns with emerging trends in teacher education research that emphasize the importance of analyzing concrete outputs to better understand teacher competence (Adipat, 2021; Haerazi, 2023; Pratolo et al., 2026). From a theoretical perspective, this study conceptualizes the quality of instructional materials as a tangible and measurable indicator of pre-service teachers' professional readiness. It establishes a link between pedagogical knowledge and its practical enactment, thereby contributing to a more nuanced understanding of teacher competence in EFL contexts. Based on the identified research gap and the need for systematic, artifact-based evaluation, this study addresses the following research questions:

1. To what extent do instructional materials developed by pre-service English teachers meet key criteria of instructional quality?
2. What strengths are demonstrated in the instructional materials developed by pre-service English teachers?
3. What challenges do pre-service English teachers face in designing instructional materials, particularly in higher-order and future-oriented competencies?

2. Method

This section outlines the methodological procedures employed to evaluate the quality of instructional materials developed by pre-service English teachers. It includes the research design, research objects and participants, the research instrument, and data collection, data analysis, and reliability procedures.

2.1. Research Design

This study employed a quantitative descriptive research design (Creswell, 2017) to systematically evaluate the quality of instructional materials developed by pre-service English teachers.

2.2. Research Objects and Participants

The objects of this study were the 28 complete instructional materials packages developed by pre-service English teachers. These materials were the outcome of a materials development workshop and had been revised in accordance with initial expert suggestions. The evaluators consisted of ten experts in English language teaching (ELT) and instructional materials development. The experts were purposively selected based on their academic qualifications, expertise and professional experience, ensuring informed and credible judgments.

2.3. Research Instrument

The primary instrument used in this study was a structured evaluation rubric. The rubric was modified from the SQRAR (Survey, Question, Read, Analyze, Recommend) framework (Sukarno, 2024) and developed to assess the instructional quality of teaching materials across ten dimensions. The ten dimensions include: 1) Curriculum Alignment, 2) Typography & Layout, 3) Language and Content Suitability, 4) Industry 4.0 Principles, 5) Society 5.0 Principles, 6) Literacy, 7) Character, 8) Culture & Local Wisdom, 9) Higher Order Thinking Skills (HOTS), and 10) 21st Century Skills (4Cs). Each dimension was rated using a 10-point Likert-type scale (1 = very poor, 10 = excellent), ensuring consistency across all evaluations and allowing greater sensitivity in distinguishing the quality of materials.

2.4. Data Collection and Analysis Procedure

The data collection was conducted in an orderly and systematic steps. The first step involved compiling and organising the 28 revised instructional material packages. The packages were then distributed to the ten expert evaluators, together with the evaluation rubric. Each expert assigned scores (1–10) for each of the ten dimensions for every material. These procedures ensured consistency, independence of judgment, and comparability of scores across evaluators.

The data were analyzed using descriptive statistical techniques, addressing the study's objective of describing material quality. Specifically: (1) The mean score was calculated for each dimension across all materials and evaluators, (2) The overall mean score was computed to represent the general quality of the instructional materials, (3) The standard deviation was calculated to examine the variability and

agreement among expert ratings. These statistical procedures provided a clear and systematic profile of strengths and weaknesses across the evaluated dimensions.

2.5. Trustworthy

To ensure methodological rigor, both content validity and inter-rater reliability were established.

a) Content validity

The evaluation rubric was adapted from the established SQAR framework, which is aligned with the principles of instructional material evaluation. Content validity was further strengthened through expert validation prior to data collection, ensuring that the rubric was relevant, clear, and appropriate for the evaluation context.

b) Inter-rater reliability

To ensure consistency among evaluators, inter-rater reliability was assessed using the Intraclass Correlation Coefficient (ICC), which is appropriate for measuring agreement among multiple raters evaluating the same set of objects. All evaluators were provided with standardized guidelines prior to scoring to minimize subjectivity and enhance scoring consistency. The resulting ICC value indicates the degree of agreement among raters, thereby providing statistical evidence of the reliability of the evaluation results.

3. Findings and Discussion

3.1. Findings

This section includes a detailed description and prescriptive evaluation of an expert review of 28 instructional materials created by pre-service teachers of English as a foreign language. This review was conducted in conjunction with other educational theories and studies to provide a thorough explanation of participants' post-workshop competence in instructional design.

Table 1. Evaluation Summary of Instructional Materials Developed by Pre-service English Teachers

No	Aspect	Score (1-10 for each criterion)										Total score
		Curriculum	Typography	Suitability	IR 4.0	Society 5.0	Literacy	Character Edu	Culture & Local Wisdom	HO TS	21st C skills	
1	Sum	220	213	215	196	196	214	213	213	195	206	2081
2	Average	7.857	7.607	7.679	7.000	7.000	7.643	7.607	7.607	6.964	7.357	74.321
3	Std	0.770	1.121	0.898	1.095	1.059	0.730	0.934	1.340	0.934	0.884	5.844
4	Modus	8	8	8	7	7	8	8	8	7	8	80
5	Median	8	8	8	7	7	8	8	8	7	8	79

Table 1 presents the descriptive statistical results of expert evaluations of the 28 instructional material packages developed by pre-service English teachers. These materials, produced as the outcome of a structured workshop, represent the participants' ability to integrate pedagogical principles into practical instructional design. The evaluation was conducted by ten experts in English language teaching and materials development, generating a comprehensive dataset consisting of mean, standard deviation, median, and mode for each of the ten assessed dimensions.

The instructional materials achieved an average total score of 74.321 out of a possible 100, indicating a generally moderate to high level of quality. This suggests that the pre-service teachers were able to demonstrate an adequate level of competence in developing instructional materials, although there remains room for further refinement. The relatively close clustering of central tendency measures (mean, median, and mode) across most dimensions indicates that the evaluations were consistently centred within the upper-mid range, reflecting a stable perception of material quality among the expert raters.

A closer examination of the mean scores reveals a clear pattern of strengths and weaknesses across dimensions. Curriculum alignment emerged as the strongest aspect ($M = 7.857$), indicating that the

instructional materials were generally well-aligned with learning objectives and curricular standards. This finding suggests that the pre-service teachers possess a solid understanding of curriculum requirements and are capable of translating them into instructional content. Similarly, relatively high scores in suitability ($M = 7.679$), literacy ($M = 7.643$), typography ($M = 7.607$), character education ($M = 7.607$), and culture and local wisdom ($M = 7.607$) indicate that the materials were generally appropriate for learners, linguistically accessible, and culturally contextualized. These results imply that the participants were able to address fundamental pedagogical and contextual considerations effectively.

In contrast, lower mean scores were observed in higher-order thinking skills (HOTS) ($M = 6.964$), Industry 4.0 ($M = 7.000$), Society 5.0 ($M = 7.000$), and 21st-century skills ($M = 7.357$). These findings are particularly significant, as they highlight a gap between foundational instructional design competencies and more advanced, future-oriented pedagogical integration. The relatively lower performance in HOTS suggests that while the materials may support basic comprehension and application, they are less effective in promoting critical thinking, analysis, and problem-solving. Similarly, the modest scores in Industry 4.0 and Society 5.0 dimensions indicate that the integration of technology-driven, innovation-oriented, and human-centred competencies remains limited. This pattern suggests that pre-service teachers may require more explicit training in embedding these contemporary educational paradigms into instructional materials.

The analysis of median and mode values further reinforces these trends. Most dimensions show median and mode values of 7 or 8, indicating a consistent concentration of scores within a narrow range. Higher central tendency values in curriculum alignment, suitability, literacy, and character education confirm the relative strength of these areas, whereas the slightly lower values in HOTS, Industry 4.0, and Society 5.0 reflect comparatively weaker performance. This consistency between mean, median, and mode suggests that the observed patterns are not influenced by extreme scores but represent stable evaluation outcomes.

The standard deviation results provide additional insight into the level of agreement among expert evaluators. Lower standard deviation values in literacy ($SD = 0.730$) and curriculum alignment ($SD = 0.770$) indicate a high level of agreement, suggesting that these aspects were consistently perceived as strong across materials. In contrast, higher variability in typography ($SD = 1.121$), industry 4.0 ($SD = 1.095$), society 5.0 ($SD = 1.059$), and particularly culture and local wisdom ($SD = 1.340$) indicate less consensus among evaluators. This variability may reflect differences in expert interpretation, especially in dimensions that are more subjective or context-dependent, such as cultural representation and technological integration. It may also suggest inconsistency in how these elements were implemented across the instructional materials.

Taken together, these findings reveal a meaningful pattern: while pre-service teachers demonstrate competence in core instructional design principles, such as alignment, appropriateness, and literacy integration, they exhibit relatively weaker performance in higher-level and future-oriented competencies, including critical thinking development and integration of Industry 4.0 and Society 5.0 principles. This indicates a need for more targeted pedagogical support that emphasizes not only foundational skills but also the development of innovative, critical, and technology-integrated instructional practices. Such an imbalance highlights an important implication for teacher education programs, suggesting that curriculum and training should place greater emphasis on bridging the gap between traditional instructional design and the demands of contemporary education.

a) Strengths

In response to Research Question 2, which examines the strengths demonstrated in instructional materials developed by pre-service English teachers, the results reveal several areas of competence, particularly in curriculum alignment, pedagogical appropriateness, and the application of established instructional frameworks. Curriculum alignment emerged as the strongest dimension, with the highest mean score ($M = 7.857$). This indicates that, overall, participants were able to design instructional materials that correspond with prescribed curriculum objectives. This strength is further reflected in the relative consistency of scores within this dimension, suggesting that most participants demonstrated at least a baseline ability to identify and incorporate relevant learning goals into their materials. However, the fact that the score does not reach a higher threshold (e.g., above 8.5) implies that this alignment may not be systematically or comprehensively achieved. In practical terms, while learning objectives are generally present and identifiable, their connection to specific activities,

instructions, or expected outcomes may vary in clarity and coherence across materials. This pattern suggests that participants are able to recognize curricular demands but may not consistently operationalize them into fully integrated instructional designs.

A similar pattern can be observed in the dimension of suitability ($M = 7.679$), which reflects the degree to which materials are appropriate for learners' proficiency levels and learning contexts. The relatively high mean score indicates that participants were generally successful in adjusting language level, task difficulty, and content selection to match target learners. This suggests an emerging awareness of learner characteristics and instructional needs. The literacy dimension ($M = 7.643$) provides further evidence of participants' ability to produce materials with clear and comprehensible language. The results suggest that instructions, explanations, and textual elements within the materials are generally understandable and accessible to learners. This reflects a functional level of language control and awareness of clarity in instructional communication.

In terms of typography ($M = 7.607$), the findings indicate that participants were able to apply basic principles of visual organization and layout. This includes the use of readable fonts, appropriate spacing, and general structuring of content to support usability. While this suggests an understanding of the importance of visual clarity in instructional materials, the score also implies that more advanced design considerations, such as visual hierarchy, emphasis, and integration of multimodal elements, may not be fully developed. The 21st-century skills (4Cs) dimension yielded a slightly lower, yet still relatively strong, mean score ($M = 7.357$), indicating that participants were able to incorporate elements of communication, collaboration, critical thinking, and creativity into their instructional materials. This is evident in the inclusion of commonly used classroom activities such as group discussions, presentations, and problem-solving tasks. When considered collectively, the pattern of findings reveals that pre-service teachers demonstrate a cluster of strengths centered on foundational and procedural aspects of instructional design. The relatively close range of mean scores (7.357–7.857) across all dimensions suggests a balanced but moderate level of competence rather than the presence of highly advanced expertise in any single area. This consistency indicates that participants have acquired a shared baseline of knowledge and skills in material development.

b) Challenges

In response to Research Question 3, which examines the challenges faced by pre-service English teachers in designing instructional materials, particularly in relation to higher-order and future-oriented competencies, the findings reveal consistent limitations in several key dimensions. Specifically, higher-order thinking skills (HOTS) recorded the lowest mean score ($M = 6.964$), followed by Industry 4.0 ($M = 7.000$) and Society 5.0 ($M = 7.000$), indicating that these areas were less successfully integrated compared to other aspects of instructional design. These lower mean scores suggest that participants encountered difficulties in embedding complex cognitive demands and contemporary educational paradigms into their materials, pointing to a gap between foundational instructional competence and more advanced pedagogical expectations.

This pattern is further supported by the variability observed in expert evaluations. The standard deviation values for Industry 4.0 ($SD = 1.095$) and Society 5.0 ($SD = 1.059$) indicate a moderate level of inconsistency across instructional materials, suggesting that participants differed in their ability to interpret and apply these concepts. Although HOTS demonstrated slightly lower variability ($SD = 0.934$), it still reflects noticeable differences in how effectively higher-order thinking components were incorporated. Such variation implies that while some participants were able to partially integrate these dimensions, others struggled more significantly, resulting in uneven quality across the dataset.

The analysis of central tendency measures further reinforces these findings. The median and mode values for HOTS, Industry 4.0, and Society 5.0 consistently remained at 7, whereas higher-performing dimensions such as curriculum alignment, suitability, and literacy reached values of 8. This indicates that the overall distribution of scores for higher-order and future-oriented competencies was systematically lower, rather than being influenced by isolated cases. In other words, the challenge appears to be structural and widespread among participants, rather than incidental or limited to a small number of individuals.

A similar trend is evident in the distribution of total scores across instructional materials. While several dimensions achieved relatively stable and higher score ranges, the scores for HOTS, Industry 4.0, and Society 5.0 were more concentrated in the mid-range of the scale. Notably, these dimensions less frequently reached the upper end of the scoring range, indicating that only a limited number of

instructional materials demonstrated strong performance in these areas. The majority of materials remained at a moderate level, suggesting that participants were generally able to meet minimum expectations but rarely achieved advanced levels of integration.

Taken together, the combination of lower mean scores, lower central tendency values, and moderate variability indicates that higher-order and future-oriented competencies represent a significant challenge for pre-service teachers. These findings suggest that participants may lack sufficient pedagogical knowledge and practical experience in designing instructional activities that promote critical thinking, technological integration, and innovation-oriented learning. More importantly, the results point to a deeper issue in instructional design development, namely the difficulty of transforming abstract educational concepts, such as HOTS, Industry 4.0, and Society 5.0, into concrete, implementable classroom practices. This highlights the need for more targeted instructional support within teacher education programs, particularly in bridging the gap between theoretical understanding and practical application of advanced pedagogical competencies.

3.2 Discussion

This study set out to examine the quality of instructional materials developed by pre-service English teachers, with particular attention to their strengths and the challenges they face in integrating higher-order and future-oriented competencies. Taken together, the findings reveal a consistent and meaningful pattern: while the instructional materials demonstrate a moderate to high level of overall quality, this competence is unevenly distributed, with clear strengths in foundational aspects of instructional design and notable limitations in more complex, cognitively demanding, and future-oriented dimensions. With regard to overall quality (RQ1), the average score of 74.321 indicates that the materials generally meet acceptable standards of instructional design. This suggests that pre-service teachers are able to apply core pedagogical principles, particularly in organizing content, aligning materials with curricular goals, and ensuring basic usability. The relatively tight clustering of mean, median, and mode across dimensions further indicates that this level of performance is stable and consistently recognized by expert evaluators. From a theoretical perspective, this reflects successful acquisition of foundational instructional competencies, which, as argued Mukundan et al. (2011), are grounded in the alignment between objectives, content, and learning experiences. In this sense, the findings suggest that teacher education programs are effective in developing baseline instructional design skills.

However, a more nuanced analysis of the findings reveals that this overall competence is largely concentrated in foundational and procedural domains, as highlighted in RQ2. The strongest performance in curriculum alignment, along with relatively high scores in suitability, literacy, typography, and cultural relevance, indicates that participants are designing materials that are appropriate, accessible, and contextually grounded. These strengths align with established criteria of effective materials evaluation, particularly those proposed by Laabidi (2021) who emphasize the importance of relevance to learners' needs and alignment with curricular objectives. At this level, pre-service teachers demonstrate the ability to produce materials that are pedagogically functional and responsive to immediate instructional requirements (Ann, & Castañeda, 2025; Drajadi et al., 2021; Limbong et al., 2024).

Nevertheless, these strengths appear to be predominantly procedural rather than deeply conceptual. While participants can successfully incorporate familiar pedagogical elements, such as structured activities, group work, and basic communicative tasks, the findings suggest that these are often implemented as standard formats rather than as part of a coherent instructional design strategy. This distinction is critical. As highlighted by Basikin, (2024), effective instructional design requires not only identifying learning objectives but also ensuring that all components of instruction, tasks, assessments, and learning experiences, are intentionally aligned to promote deep understanding. The present findings suggest that while pre-service teachers understand "what to teach," they are less consistent in designing how learning should meaningfully occur (Garba et al., 2013; Legesse et al., 2024; Pratiwi & Waluyo, 2022). This indicates an emerging but incomplete mastery of instructional design thinking (Kao, 2023)

The challenges identified in RQ3 further illuminate this gap. The consistently lower performance in higher-order thinking skills (HOTS), Industry 4.0, and Society 5.0 dimensions suggests that participants encounter significant difficulty when engaging with higher-level cognitive demands and contemporary educational paradigms. From a cognitive standpoint, this reflects limited engagement

with the upper levels of Benjamin Bloom's taxonomy, particularly in promoting analysis, evaluation, and creation. The instructional materials appear to prioritize comprehension and application, but less frequently extend to tasks that require critical inquiry, problem-solving, or knowledge construction (Laabidi, 2021; Setiyowati et al., 2025).

In addition, the relatively modest integration of Industry 4.0 and Society 5.0 principles indicates that the incorporation of technological, innovative, and human-centered competencies remains underdeveloped. This finding is particularly significant in light of contemporary educational demands, where learners are expected to engage with complex, technology-mediated environments (Bueno et al., 2023; Erfan et al., 2022; Joshi, 2023). The Technological Pedagogical Content Knowledge (TPACK) framework proposed by Das and Mitra, (2021); Zhang et al. (2025) provides a useful lens for interpreting this limitation. According to Namestovski & Kovari (2022) framework, effective teaching requires the integration of content, pedagogy, and technology (Garba et al., 2013; Legesse et al., 2024; Olori & Saka, 2024). However, the current findings suggest that pre-service teachers' expertise remains more strongly developed in content and pedagogy than in technology integration and innovation-oriented design.

Importantly, the variability observed in these dimensions indicates that these challenges are not uniformly distributed but reflect differing levels of understanding and implementation among participants. This suggests that competencies related to HOTS and future-oriented frameworks are not yet systematically internalized within the teacher education process (Zhang et al., 2025). From a sociocultural perspective, as argued by Lev Vygotsky, such differences may arise from variations in learning experiences, scaffolding, and practice opportunities (Limbong et al., 2024). However, the overall pattern indicates that these competencies remain developmental rather than established (Greene et al., 2023; Grewal et al., 2024).

Taken together, the findings point to a critical imbalance in the instructional design capabilities of pre-service teachers. Their strengths lie in what may be described as the "visible" dimensions of teaching, those that are easier to operationalize, such as alignment, organization, and basic activity design. In contrast, their challenges are situated in the "invisible" yet cognitively demanding dimensions, including critical thinking development, technological integration, and innovation-oriented pedagogy. This distinction is particularly important because it highlights a gap between surface-level instructional competence and deeper pedagogical expertise. This imbalance carries important implications for teacher education. While current programs appear effective in building foundational skills, they may place insufficient emphasis on developing higher-order instructional design thinking and the ability to translate abstract educational concepts into concrete classroom practices.

4. Conclusion

This study investigated the quality of instructional materials developed by pre-service English teachers, focusing on overall performance, demonstrated strengths, and challenges in integrating higher-order and future-oriented competencies. The findings indicate that the instructional materials generally achieved a moderate to high level of quality, suggesting that pre-service teachers possess adequate foundational skills in instructional design. In particular, they demonstrated strength in curriculum alignment, learner suitability, and the application of familiar pedagogical frameworks, indicating their ability to produce materials that are relevant, organized, and pedagogically functional. However, the study also reveals a critical imbalance in the distribution of these competencies. While participants performed well in foundational and procedural aspects of instructional design, they encountered notable challenges in integrating higher-order thinking skills (HOTS), as well as Industry 4.0 and Society 5.0 principles. These dimensions, which represent more complex, cognitively demanding, and future-oriented competencies, were less consistently and less effectively implemented. This suggests that pre-service teachers may understand instructional objectives at a conceptual level but face difficulties in translating them into meaningful, cognitively engaging, and technologically enriched learning experiences.

The findings highlight an important gap between basic instructional competence and advanced pedagogical expertise. More specifically, pre-service teachers appear to be proficient in designing the structural and visible components of instruction but less capable of embedding deeper cognitive processes and innovative elements within their materials. This gap underscores the need to

reconceptualize teacher preparation, not only as the development of technical skills but also as the cultivation of higher-order instructional thinking and design capacity. From a pedagogical perspective, these findings suggest that teacher education programs should place greater emphasis on explicitly teaching how to design tasks that promote critical thinking, problem-solving, and knowledge construction. This includes providing structured opportunities for pre-service teachers to practice designing HOTS-oriented activities and to receive targeted feedback on their instructional decisions. Furthermore, the integration of Industry 4.0 and Society 5.0 principles should be more systematically embedded within the curriculum, not merely as theoretical concepts but as practical design challenges that require the application of technology, innovation, and real-world problem-solving.

Subsequent investigations should focus on a few key areas. It is crucial to focus on intervention studies that construct and evaluate models of professional learning (workshop plus coaching) on the identified weak areas. Longitudinal studies that examine the growth of early-career novice teachers' skill acquisition and development in the area of complexity will also be important. Ultimately, research is necessary to develop constructivist frameworks that enable teachers to effectively implement broad educational policies, such as Society 5.0, within the classroom context.

Acknowledgment

The authors express their profound gratitude to all individuals who played a role in the research undertaking, including those involved in the collection and examination of data, as well as those who provided constructive feedback.

Declarations

- Author contribution** : All authors contributed to the research process and manuscript preparation. Sukarno, Sari Hidayati, Ihtiar Fitrianiingsih, Dewi Nur Widiyati, and Shabrina Yumna Azhra contributed to the conceptualization of the study, research design, instrument adaptation, data analysis, interpretation of findings, and manuscript writing. Previously, Sukarno had developed the instrument. Megan Jinabe, Fitriya Dessi Wulandari, Herwin Dyastuti, Kauna Karima, Adivta Yudha Tama, and Nita Maghfiratul Jannah assisted in data collection, data organization, and manuscript preparation. Fadhl Mohammed Awadh Gharamah and Yasir A. Alsamiri assisted in data analysis, manuscript review, and academic refinement. Final revision was made by Sukarno, Sari Hidayati, Ihtiar Fitrianiingsih, Dewi Nur Widiyati, Shabrina Yumna Azhra, Megan Jinabe, and Fitriya Dessi Wulandari. All authors have read and approved the final version of the manuscript.
- Funding statement** : This research was funded by Universitas Negeri Yogyakarta.
- Conflict of interest** : There is no conflict of interest in this work
- Ethical declaration** : We as authors acknowledge that this work has been written based on ethical research that conforms with the regulations of our university and that we have obtained the permission from the relevant institute when collecting data.
- We support ELTEJ in maintaining high standards of personal conduct, practicing honesty in all our professional practices and endeavours.
- Additional information** : No additional information is available for this paper.

REFERENCES

- Adipat, S. (2021). Developing Technological Pedagogical Content Knowledge (TPACK) through Technology-Enhanced Content and Language-Integrated Learning (T-CLIL) Instruction. *Education and Information Technologies*, 26(5), 6461–6477. <https://doi.org/10.1007/s10639-021-10648-3>
- Amalia, T., Inayati, D., & Marini, A. (2023). Improving students' motivation in learning English through gamification. *Jurnal Pendidikan Bahasa Inggris Undiksha*, 11(1), 1–9. <https://doi.org/10.23887/jpbi.v11i1.63332>
- Ann L. Jaca, C., & D. Castañeda Jr., W. (2025). Graduate competencies as predictors of the pre-service English teachers' work-readiness. *International Journal of Evaluation and Research in Education (IJERE)*, 14(4), 3211. <https://doi.org/10.11591/ijere.v14i4.32609>
- Azhra, S. Y., Carabella, P. A., Eiden, K., Payung, N. F., Sukarno, S., Pratiwi, D., & Gharamah, F. M. A. (2024). Improving junior high school students' vocabulary understanding by using quiz-based application. *Diksi*, 32(2), 247–266. <https://doi.org/10.21831/diksi.v32i2.69225>
- Basikin, B. (2024). The Contribution of Inservice Teacher Education Program (PPG) on teachers' professionalism. *English Language Teaching Educational Journal*, 6(2), 124–136. <https://doi.org/10.12928/eltej.v6i2.10076>
- Bueno, R., Niess, M. L., Aldemir Engin, R., Ballejo, C. C., & Lieban, D. (2023). Technological pedagogical content knowledge: Exploring new perspectives. *Australasian Journal of Educational Technology*, 88–105. <https://doi.org/10.14742/ajet.7970>
- Bulqiyah, S., & Munir, A. (2023). Revealing EFL teachers' experiences in the integration of web-based assessment media for assessment. *Journal of English Teaching and Learning Issues*, 6(1), 45. <https://doi.org/10.21043/jetli.v6i1.19624>
- Creswell, J. W. (2017). *Qualitative inquiry and research design (international student edition). Choosing among five approaches*. SAGE Publications.
- Das, K., & Mittra, S. (2021). Scope of introducing TPACK of teacher education programme in India. *Shanlax International Journal of Arts, Science and Humanities*, 8(4), 31–38. <https://doi.org/10.34293/sijash.v8i4.3805>
- Drajati, N. A., Rakerda, H., Sulistyawati, H., Nurkamto, J., & Ilmi, M. (2021). Investigating the adoption of TPACK-21cl by English pre-service teachers in a Covid-19 teaching practicum. *Indonesian Journal of Applied Linguistics*, 11(1). <https://doi.org/10.17509/ijal.v11i1.34625>
- Erfan, F., Sulisyaningsih, S., & Syakur, Abd. (2022). Integration of ICT-pedagogy for senior high school students: Pre-service teachers' TPACK (Technological Pedagogical Content Knowledge). *Budapest International Research and Critics Institute (BIRCI-Journal)*. <https://doi.org/10.33258/birci.v5i1.4236>
- Esteban, A. J., & Sung, K. (2025). Influence of playing online video games on Filipino college students' confidence in speaking English. *International Journal of Evaluation and Research in Education (IJERE)*, 14(4), 2555. <https://doi.org/10.11591/ijere.v14i4.32842>
- Febriyanti, E. R. (2018). Identifikasi analisis kebutuhan pembelajar bahasa Inggris (non Program Studi Bahasa Inggris) pada mata kuliah bahasa Inggris ESP di lingkungan FKIP Universitas Lambung Mangkurat Banjarmasin. *Vidya Karya*, 32(2), 123. <https://doi.org/10.20527/jvk.v32i2.5230>
- Garba, S. A., Ranjit Singh, T. K., & Yusuf, N. M. (2013). Integrating technology in teacher education curriculum and pedagogical practices: The effects of web-based technology resources on pre-service teachers' achievement in teacher education training. *Proceedings of the 2013 International Conference on Information Science and Technology Applications*. <https://doi.org/10.2991/icista.2013.14>
- Gorozhankina, T. N., & Grechukhina, V. V. (2022). Authentic video materials as a means of shaping communicative competence in Russian among Spanish and Dutch learners. *Russian Language Studies*, 20(4), 483–499. <https://doi.org/10.22363/2618-8163-2022-20-4-483-499>

- Greene, M., Cheng, S.-L., & Jones, M. (2023). Preservice teachers' technology integration knowledge development in an online technology-based course. *International Journal of Instruction*, 16(4), 385–404. <https://doi.org/10.29333/iji.2023.16423a>
- Grewal, S., Manuel, M., & Veettil, R. P. (2024). Principles and procedures of material development in the evolving ELT scenario. *Salud, Ciencia y Tecnología - Serie de Conferencias*, 3, 656. <https://doi.org/10.56294/sctconf2024656>
- Haerazi, H. (2023). English pre-service teachers' responses to writing online instruction and critical teacher feedback. *Pedagogy: Journal of English Language Teaching*, 11(2), 141. <https://doi.org/10.32332/joelt.v11i2.7388>
- Hamzah, I. (2023). The effect of Quizizz on students reading comprehension. *Journal of Arts and Education*, 3(1). <https://doi.org/10.33365/jae.v3i1.208>
- Handayani, S., Hussin, M., & Norman, H. (2024). Evaluating teaching readiness using the TPACK model: Factor, reliability, and validity analyses for Indonesian Economics teacher candidates. *Perspectives of Science and Education*, 68(2), 679–698. <https://doi.org/10.32744/pse.2024.2.41>
- Hoir, S., Subandowo, M., & Wiyarno, Y. (2019). Improve your English through the 21st century skill. *Edcomtech Jurnal Kajian Teknologi Pendidikan*, 4(1), 38–48. <https://doi.org/10.17977/um039v4i12019p038>
- Husnah, L. Al, Heriyawati, D. F., & Elfiyanto, S. (2023). Quizizz paper mode is new: Students' perception of using E-Tool of language assessment in EFL class. *Celtic: A Journal of Culture, English Language Teaching, Literature and Linguistics*, 10(2), 121–137. <https://doi.org/10.22219/celtic.v10i2.28528>
- Joshi, S. (2023). TPACK and teachers' self-efficacy: A systematic review. *Canadian Journal of Learning and Technology*, 49(2), 1–23. <https://doi.org/10.21432/cjlt28280>
- Julita, R. (2024). The impact of gamification on EFL students' reading comprehension. *Journal of English as a Foreign Language Education (JEFLE)*, 4(2), 136. <https://doi.org/10.26418/jefle.v4i2.74901>
- Kao, Y. (2023). Empowering preservice English teachers with language assessment literacy concepts and practices: Application of Vygotskian concept-based language instruction. *The Modern Language Journal*, 107(S1), 68–94. <https://doi.org/10.1111/modl.12818>
- Kapoyannis, T. (2021). The Name Jar Project: Supporting preservice teachers in working with English language learners. *Language and Literacy*, 23(3), 45–65. <https://doi.org/10.20360/langandlit29510>
- Laabidi, Y. L. Y. (2021). Influence of teaching experience on teachers' level of use of critical thinking. *LLT Journal: A Journal on Language and Language Teaching*, 24(2), 514–528. <https://doi.org/10.24071/llt.v24i2.3719>
- Legesse, B., Legesse, D., & Demissie, M. (2024). Examining the mediating role of college-based capacities on the relationship between technology integration and teacher preparation program implementation quality in the Southern Nation, nationalities and peoples' region, Ethiopia. <https://doi.org/10.21203/rs.3.rs-4973025/v1>
- Limbong, E., Setiawan, I., & Hamilton, A. (2024). Bridging the gap: The reality of digital technology integration by Indonesian pre-service EFL teachers. *Script Journal: Journal of Linguistics and English Teaching*, 9(1), 58–78. <https://doi.org/10.24903/sj.v9i1.1524>
- Liu, P.-L., Chen, C.-J., & Huang, C.-Y. (2025). Evaluating the effectiveness of Quizizz for enhancing English vocabulary acquisition and engagement among rural elementary students. *Sage Open*, 15(4). <https://doi.org/10.1177/21582440251379697>
- Maming, K., Sani, A. A., Putra, D. P., & Radiana, S. P. (2023). Get ahead with Quizizz: Advancing junior high school students' vocabulary mastery in online learning. *Elsya: Journal of English Language Studies*, 5(2), 224–235. <https://doi.org/10.31849/elsya.v5i2.10275>
- Manipatruni, V. R., Kumar, N. S., Karim, M. R., & Banu, S. (2023). Improving English vocabulary through Quizizz in practice tests for gamification and Google Forms with AutoProctor in

- assessment tests for the preclusion of malpractice. *International Journal of Interactive Mobile Technologies (IJIM)*, 17(13), 22–43. <https://doi.org/10.3991/ijim.v17i13.39445>
- Mat, H., Nusantara, T., Atmoko, A., Hanafi, Y., & Mustakim, S. S. (2025). Need analysis: development of a teaching module for enhancing higher-order thinking skills of primary school students. *International Journal of Evaluation and Research in Education (IJERE)*, 14(3), 1643. <https://doi.org/10.11591/ijere.v14i3.30335>
- Mukundan, J., Hajimohammadi, R., & Nimehchisalem, V. (2011). Developing an English language textbook evaluation checklist. *Contemporary Issues in Education Research (CIER)*, 4(6), 21. <https://doi.org/10.19030/cier.v4i6.4383>
- Namestovski, Ž., & Kovari, A. (2022). Framework for preparation of engaging online educational materials, a cognitive approach. *Applied Sciences*, 12(3), 1745. <https://doi.org/10.3390/app12031745>
- Nugraheni, R., & Fakhurriana, R. (2023). Boosting English vocabulary proficiency in EFL learners: A study on the efficacy of Quizizz at IAIN Kediri. *Indonesian Journal of Multidisciplinary Educational Research*, 1(2), 161–177. <https://doi.org/10.30762/ijomer.v1i2.1379>
- Obasi, J. C. (2022). “What I see, I understand”: Attitudes of selected university undergraduates towards the utilization of authentic materials in teaching in an ESL classroom. *Nsukka Journal of the Humanities*, 30(1), 21–31. <https://doi.org/10.62250/nsuk.2022.30.1.21-31>
- Ölçer-Çevik, S., & Kozaner, Ç. (2025). Building a sustainable future: perspectives on primary school sustainability projects from teachers and pre-service teachers. *International Journal of Evaluation and Research in Education (IJERE)*, 14(5), 3403. <https://doi.org/10.11591/ijere.v14i5.33362>
- Olori, A., & Saka, A. O. (2024). Differences in Technological Pedagogical Content Knowledge between experienced and less experienced computer teachers in Ogun State, Nigeria. *IJUM Journal of Educational Studies*, 12(2), 21–32. <https://doi.org/10.31436/ijes.v12i2.514>
- Permana, P., Permatawati, I., & Khoerudin, E. (2023). Foreign language learning gamification using Quizizz: A systematic review based on students’ perception. *Eralingua: Jurnal Pendidikan Bahasa Asing dan Sastra*, 7(2), 233. <https://doi.org/10.26858/eralingua.v7i2.23969>
- Pratolo, B. W., Anwari, H., Purwanti, E., Fatimah, N., Soviyah, S., & Nuryana, Z. (2026). Revolution or ruin? Systematic insights into the redefinition of the teacher book. *English Language Teaching Educational Journal*, 8(3), 249–270. <https://doi.org/10.12928/eltej.v8i3.15138>
- Pratiwi, D. I., & Waluyo, B. (2022). Integrating task- and game-based learning into an online TOEFL preparation course during the COVID-19 outbreak. *Malaysian Journal of Learning and Instruction*, 19. <https://doi.org/10.32890/mjli2022.19.2.2>
- Ridwan, A., & Athena, T. (2023). Professional development altogether: Pre-service and in-service EFL teacher collaboration in doing classroom action research. *Elsya: Journal of English Language Studies*, 5(2), 147–158. <https://doi.org/10.31849/elsya.v5i2.11929>
- Setiyowati, E., Puspita, A. M. I., & Mariana, N. (2025). Teachers’ pedagogical competence in designing Deep Learning and HOTS assessment in elementary schools. *Journal of Innovation and Research in Primary Education*, 4(4), 4011–4029. <https://doi.org/10.56916/jirpe.v4i4.2633>
- Setyaningrum, A., & Putri, D. (2023). Pre-service English teachers’ voices toward the use of Quizizz in learning tourism. *Pedagogy: Journal of English Language Teaching*, 11(2), 163. <https://doi.org/10.32332/joelt.v11i2.7629>
- Stofflett, R. T., & Stoddart, T. (1994). The ability to understand and use conceptual change pedagogy as a function of prior content learning experience. *Journal of Research in Science Teaching*, 31(1), 31–51. <https://doi.org/10.1002/tea.3660310105>
- Sukarno, S. (2024). SQRAR: A model of coursebook evaluation. *English Language Teaching Educational Journal*, 7(1), 13–23. <https://doi.org/10.12928/eltej.v7i1.10306>

- Sumarni, S., Ifan Iskandar, & Imam Santosa. (2023). Assessing digital competence among pre-service English language teachers: Strengths and weaknesses. *Journal of English Education and Teaching*, 7(3), 613–631. <https://doi.org/10.33369/jeet.7.3.613-631>
- Thohir, M. A., Sukarelawan, Moh. I., Jumadi, J., Warsono, W., & Citrasukmawati, A. (2021). The effects of an instructional design-based web course on pre-service teachers' competencies. *International Journal of Evaluation and Research in Education (IJERE)*, 10(1), 230. <https://doi.org/10.11591/ijere.v10i1.20475>
- Tortorelli, L. S., Lupo, S. M., & Wheatley, B. C. (2021). Examining teacher preparation for code-related reading instruction: An integrated literature review. *Reading Research Quarterly*, 56(S1). <https://doi.org/10.1002/rrq.396>
- Waluyo, B., Phanrangsee, S., & Whanchit, W. (2023). Gamified grammar learning in online English courses in Thai higher education. *Online Journal of Communication and Media Technologies*, 13(4), e202354. <https://doi.org/10.30935/ojcm/13752>
- Zein, S., Sukyadi, D., Hamied, F. A., & Lengkanawati, N. S. (2020). English language education in Indonesia: A review of research (2011–2019). *Language Teaching*, 53(4), 491–523. <https://doi.org/10.1017/S0261444820000208>
- Zhang, H., Wang, Z., Jiang, R., & Wu, H. (2025). Exploring the influence of technology-enhanced active learning environments on pre-service teachers' TPACK and technology beliefs. *Sage Open*, 15(3). <https://doi.org/10.1177/21582440251359823>