



Reconstructing a Junior High School English Question Bank through Qualitative Item Analysis and Higher-Order Thinking Skills Integration

Tedi Rohadi¹, Jamaludin², Nana Priajana³, Shofia Mardhiyah⁴

^{1,4}Islamic State University of Sunan Gunung Djati Bandung

^{2,3}Islamic State University of Cyber Syekh Nurjati Cirebon

Article Info

Received: 2026-03-17

Revised: 2026-03-29

Accepted: 2026-04-02

Keywords:

assessment design,
formative assessment,
higher-order thinking skills,
question bank,
Quizizz.

DOI:

10.24256/ideas.v14i1.9967

Corresponding Author:

Tedi Rohadi

tedi.rohadi@uinsgd.ac.id

Islamic State University of Sunan
Gunung Djati Bandung

Abstract

Assessment plays an important role in supporting learning and evaluating students' cognitive development. However, many classroom assessments still emphasize lower-order thinking skills and contain weaknesses in item construction. This study aimed to analyze the shortcomings of existing English question banks and reconstruct them into a more cognitively balanced assessment instrument for junior high school learning. The study employed a qualitative design using document analysis of sixty multiple-choice items taken from two English question banks. The items were analyzed based on Bloom's Revised Taxonomy and item-writing guidelines focusing on cognitive levels and item construction. The results showed that the existing question banks were dominated by lower-order thinking skills, particularly at the applying level, while higher-order thinking skills were rarely represented. Several structural weaknesses were also identified, including implausible distractors, heterogeneous options, and unclear stems. Based on these findings, the question bank was reconstructed by redistributing cognitive levels and improving item construction. The reconstructed items were also prepared for digital implementation using the Quizizz platform to support formative assessment. The study concludes that systematic item analysis can improve both the cognitive quality and structural validity of assessment instruments. The reconstructed question bank provides a more balanced and effective assessment tool for junior high school English learning.

1. Introduction

Assessment plays an essential role in the teaching and learning process because it provides information about students' learning progress and supports instructional decision-making. In language education, assessment is not only used to evaluate learning outcomes but also to support learning through formative feedback and meaningful cognitive engagement. Recent research highlights that effective assessment practices should encourage deeper learning and promote higher-order thinking skills rather than focusing solely on factual recall (Panadero, Andrade, & Brookhart, 2019; Carless & Winstone, 2023). Therefore, the quality of assessment instruments becomes a critical factor in ensuring that students' learning outcomes are measured accurately and meaningfully.

In recent years, educational systems have increasingly emphasized the development of higher-order thinking skills (HOTS) as an important goal of 21st-century learning. These skills involve the ability to analyze information, evaluate ideas, and apply knowledge in different contexts. Bloom's Revised Taxonomy categorizes cognitive processes into six levels: remembering, understanding, applying, analyzing, evaluating, and creating (Anderson & Krathwohl, 2001). Assessments that incorporate these cognitive levels can promote deeper learning and critical thinking development. However, several studies report that classroom assessments are still dominated by lower-order thinking tasks such as remembering and understanding, while higher cognitive processes remain underrepresented (Brookhart, 2019; Tarrant & Ware, 2012). This imbalance indicates that many assessment practices have not yet fully supported the development of higher-order thinking skills.

One widely used instrument in classroom assessment is the multiple-choice question bank. Question banks allow teachers to organize and store assessment items systematically so that they can be reused for different testing purposes. When constructed properly, multiple-choice items can measure a range of cognitive skills and provide reliable indicators of student learning outcomes (Haladyna & Rodriguez, 2019). Nevertheless, the quality of multiple-choice assessments depends heavily on adherence to item-writing principles, including clear stems, plausible distractors, and homogeneous answer options. Previous studies have shown that poorly designed items may contain structural flaws such as implausible distractors, heterogeneous options, or ambiguous stems, which can reduce the validity and reliability of the assessment instrument (Gierl, Lai, & Turner, 2020; Tarrant & Ware, 2012).

The weaknesses of existing question banks may be influenced by several interconnected factors. Many assessment items focus mainly on lower cognitive levels, while higher-order thinking skills are rarely integrated. In addition, structural problems such as clue-giving options, unequal distractors, and unclear stems can affect the accuracy of assessment results. Language-related issues, including grammatical inconsistencies or ambiguous wording, may also reduce

item clarity. These problems highlight the importance of conducting systematic item analysis in order to improve both the cognitive level and structural quality of assessment instruments.

Despite the growing emphasis on HOTS and the increasing use of digital assessment platforms, existing classroom question banks remain cognitively imbalanced and structurally weak. Most previous studies tend to focus either on the development of HOTS-oriented assessment instruments or on the effectiveness of digital quiz platforms in increasing engagement. However, limited attention has been given to how existing question banks can be systematically analyzed, diagnostically evaluated, and reconstructed before being implemented in digital environments. This indicates a clear gap between theoretical assessment principles and the practical quality of question banks used in schools.

Recent technological developments have also opened new opportunities for improving classroom assessment practices. Digital quiz platforms such as Quizizz have been widely used to support formative assessment because they allow teachers to deliver quizzes interactively and provide immediate feedback to students. Research indicates that gamified digital assessments can increase student engagement, motivation, and participation during learning activities (Dichev & Dicheva, 2020; Razali et al., 2020).

Furthermore, studies on retrieval-based learning suggest that frequent formative quizzes can strengthen memory retention and enhance learning outcomes through the testing effect (Roediger & Butler, 2019). However, the effectiveness of digital assessment platforms largely depends on the quality of the question bank used within the system.

Although previous studies have investigated HOTS-based assessment and digital quiz platforms, limited research has examined how existing question banks can be systematically analyzed and reconstructed to improve both their cognitive quality and structural validity. Most studies focus either on the development of HOTS-oriented assessment instruments or on the impact of gamified quiz platforms on student engagement. Few studies have combined these approaches by analyzing existing question banks, reconstructing them based on item-writing principles and cognitive frameworks, and implementing the improved items within digital formative assessment platforms. In many cases, existing question banks are cognitively imbalanced and structurally weak, containing items that overemphasize lower-order thinking skills and violate fundamental item-writing principles.

Responding to this gap, the present study aims to: (1) analyze the cognitive distribution of existing question banks, (2) identify weaknesses in item construction using established item-writing guidelines, (3) reconstruct the question bank to achieve better cognitive balance and structural validity, and (4) integrate the reconstructed items into a digital formative assessment environment using Quizizz. The novelty of this study lies in the integration of qualitative item

analysis, HOTS-oriented reconstruction, and digital assessment implementation within a single systematic framework.

2. Method

This study employed a qualitative design using document analysis to examine and reconstruct English question banks used in junior high school assessment. The two question banks were selected purposively because they are widely used as assessment references for eighth-grade English learning in junior high schools and represent typical teacher-developed and commercially available question collections. The study aimed to identify weaknesses in existing question banks and develop a reconstructed question bank that integrates higher-order thinking skills and improved item construction. Qualitative document analysis was used to systematically evaluate assessment items and identify patterns of weaknesses in test design (Bowen, 2009).

The data consisted of two English question banks used for eighth-grade students, comprising a total of sixty multiple-choice items. The items were analyzed using an item-analysis checklist covering three aspects: content, item construction, and language clarity. The cognitive level of each item was classified based on Bloom's Revised Taxonomy, which includes remembering, understanding, applying, analyzing, evaluating, and creating (Anderson & Krathwohl, 2001).

The checklist was adapted from established multiple-choice item-writing guidelines to evaluate the structural quality of the test items (Haladyna, Downing, & Rodriguez, 2002). The checklist consisted of 26 criteria derived from established multiple-choice item-writing principles, covering three dimensions: (a) content relevance, (b) item construction quality, and (c) language clarity. Examples of the indicators include the presence of plausible distractors, homogeneous grammatical structure of options, clarity of the stem, avoidance of clue-giving options, and logical ordering of answer choices.

The analysis followed four systematic stages: (1) categorizing each item according to Bloom's Revised Taxonomy, (2) identifying item construction flaws using the checklist, (3) coding and grouping recurring weaknesses, and (4) designing reconstruction strategies based on the diagnostic findings. Data analysis was conducted through qualitative descriptive analysis. Each item was first categorized according to its cognitive level, and then evaluated based on the item-writing checklist to identify structural weaknesses such as implausible distractors, heterogeneous answer options, and unclear stems.

The results of the analysis were used as the basis for reconstructing the question bank. The reconstruction process involved redistributing the cognitive levels of the items to increase the representation of higher-order thinking skills and improving item construction by refining distractors, clarifying stems, and standardizing option structures. The reconstruction process followed specific criteria, including revising stems for clarity, redesigning distractors to ensure

plausibility, standardizing grammatical structures of options, and redesigning items to target higher cognitive processes such as inference, comparison, and contextual interpretation.

The reconstructed items were designed using a five-option multiple-choice format and prepared for implementation through the Quizizz platform to support digital formative assessment (Razali et al., 2020). To ensure trustworthiness, the reconstruction results were reviewed through expert judgment by two English education lecturers and peer checking to ensure theoretical alignment with Bloom’s taxonomy and item-writing standards.

3. Result

This study aimed to reconstruct a junior high school English question bank through qualitative item analysis and the integration of higher-order thinking skills (HOTS). The results are presented in three main sections: (1) cognitive-level distribution of the existing question banks, (2) construction weaknesses identified through item analysis, and (3) improvements in the reconstructed question bank.

Cognitive-Level Distribution of Existing Question Banks

The qualitative analysis was conducted on 60 multiple-choice items obtained from two English question banks used in junior high school learning materials. The cognitive levels of the items were classified based on the Revised Bloom’s Taxonomy. The distribution of the cognitive levels is shown in Table 1.

Table 1. Cognitive Distribution of Existing Question Banks (n = 60)

Cognitive Level	Number of Items	Percentage
C1 (Remembering)	4	6.6%
C2 (Understanding)	16	26.7%
C3 (Applying)	34	56.7%
C4 (Analyzing)	6	10%
C5 (Evaluating)	0	0%
C6 (Creating)	0	0%

These findings confirm the dominance of lower-order thinking skills in the existing question banks. As shown in Table 1, the cognitive distribution was dominated by lower-order thinking skills (LOTS). More than half of the items (56.7%) were categorized at the applying level (C3), while understanding (C2) accounted for 26.7% and remembering (C1) for 6.6%. Only a small proportion of the items (10%) reached the analyzing level (C4). Furthermore, no items were categorized as evaluating (C5) or creating (C6).

This imbalance indicates that the existing question banks primarily assessed procedural application rather than analytical reasoning. The absence of higher-level cognitive processes suggests that students were rarely required to

interpret information critically or evaluate textual meaning. Consequently, the cognitive demand of the existing items was limited in promoting higher-order thinking skills.

Construction Weaknesses in Existing Items

In addition to cognitive imbalance, qualitative item analysis based on 26 item-writing guidelines revealed several recurring construction weaknesses. These guidelines covered three main aspects: content and material, item construction, and language and grammar. The major structural weaknesses identified in the analyzed items are summarized in Table 2.

Table 2. Construction Weaknesses Identified in Existing Items

Identified Weakness	Frequency	Percentage
Clue-giving options	22	36.7%
Heterogeneous option structure	22	36.7%
Unequal option length	17	28.3%
Lack of logical/numerical order	20	33.3%
Implausible distractors	14	23.3%
Unclear or unfocused stems	8	13.3%

The table highlights recurring structural flaws that potentially reduce item validity and discrimination power. Table 2 showed that a considerable proportion of items did not comply with established item-writing principles. The most frequent weakness was the presence of clue-giving options, which appeared in 36.7% of the items. These clues were often caused by heterogeneous grammatical structures or implausible distractors that allowed students to identify the correct answer without fully understanding the stimulus.

Additionally, 33.3% of the items did not arrange the options logically or numerically, while 28.3% had options with unequal lengths. Such inconsistencies may unintentionally guide test-takers toward the correct answer through superficial patterns rather than comprehension. Furthermore, 23.3% of the items contained implausible distractors that were easily eliminated. In several cases, the stems were unclear or unfocused, which required students to read the options first in order to understand the question. These weaknesses potentially reduced the validity and discriminative power of the items.

Reconstruction of the Question Bank

Based on the diagnostic findings, the question bank was reconstructed through systematic redesign. The reconstruction process aimed to address the identified shortcomings while integrating higher-order thinking skills into the assessment items. One of the primary improvements was the redistribution of

cognitive levels in the reconstructed question bank.

Table 3. Cognitive Redistribution Before and After Reconstruction

Cognitive Level	Existing (n = 60)	Reconstructed (n = 40)
C1 (Remembering)	4 (6.6%)	4
C2 (Understanding)	16 (26.7%)	8
C3 (Applying)	34 (56.7%)	12
C4 (Analyzing)	6 (10%)	16
C5 (Evaluating)	0	0
C6 (Creating)	0	0

The reconstruction clearly shifts the cognitive emphasis from procedural application to analytical reasoning. The number of reconstructed items was reduced to 40 because several items were considered irreparable due to severe structural and cognitive flaws, making revision less effective than redevelopment. As presented in Table 3, the reconstructed question bank significantly increased the proportion of analyzing (C4) items. In the revised version, analyzing became the dominant cognitive level, accounting for 16 of the 40 items. This redistribution indicates a stronger emphasis on higher-order thinking processes compared to the original question banks.

The redesigned items were developed to encourage analytical reasoning through inference-making, comparison of information across texts, and interpretation of contextual meaning. Rather than relying on simple recall or recognition, the items required students to identify relationships between ideas and interpret implied messages in the given texts.

Structural Improvements in the Reconstructed Items

Beyond cognitive redistribution, the reconstruction process also improved item construction and language clarity. All reconstructed items were designed to eliminate clue-giving options and ensure that distractors were plausible and grammatically homogeneous. The options were written with relatively equal length and arranged in logical order to prevent superficial guessing strategies.

Moreover, each item in the reconstructed question bank was developed with five options instead of four. The additional distractor was intended to strengthen item discrimination and reduce the probability of random guessing. All stems were revised to ensure clarity and focus so that the main idea of the question was clearly stated before the options.

The reconstructed question bank was also implemented through a digital platform for formative assessment. The computer-based format allowed randomized item delivery and immediate feedback, which helped minimize answer-pattern recognition and improved assessment functionality.

Thus, the results demonstrate that qualitative item analysis provided a systematic foundation for reconstructing a more cognitively balanced and structurally sound English question bank. The reconstructed items better align with item-writing standards and incorporate higher-order thinking skills, making them more suitable for junior high school formative assessment.

4. Discussions

The persistence of LOTS-dominated items in teacher-developed question banks may be influenced by limited training in item writing, time constraints, and reliance on traditional assessment habits. The above findings of this study reveal that the existing junior high school English question banks were dominated by lower-order thinking skills (LOTS), particularly applying (C3), while higher-order thinking skills (HOTS) such as evaluating and creating were entirely absent. This pattern aligns with previous research indicating that many school assessments emphasize lower cognitive levels.

For instance, Abosalem (2016) reported that teacher-made tests largely focus on knowledge, comprehension, and application levels, with minimal representation of analytical or evaluative tasks. Such imbalance limits opportunities for students to develop critical thinking and analytical reasoning skills, which are essential competencies in contemporary education. The results of the present study therefore confirm the broader trend that traditional assessment practices tend to prioritize procedural knowledge rather than higher-level cognitive processes.

The reconstruction of the question bank in this study addressed this imbalance by increasing the proportion of analyzing (C4) items and integrating tasks that required inference, comparison, and contextual interpretation. This approach is consistent with the theoretical framework of Bloom's Revised Taxonomy, which emphasizes the importance of progressively engaging students in higher cognitive processes. Prior studies have demonstrated that assessment design plays a crucial role in shaping cognitive engagement.

For example, Mahendra (2022) found that integrating higher-order thinking tasks within formative assessment significantly improves students' analytical and problem-solving abilities. Similarly, Rizki et al. (2022) reported that the use of Quizizz-based assessments enhanced students' higher-order thinking skills compared with other gamified assessment platforms. These findings support the argument that assessment structures must intentionally incorporate HOTS-oriented items to promote deeper cognitive engagement.

Another key contribution of this study lies in the systematic improvement of item construction based on established item-writing guidelines. The qualitative item analysis identified several structural issues in the existing question banks, including clue-giving options, implausible distractors, and inconsistent option structures. These issues can compromise item validity by allowing students to

guess the correct answer through superficial cues rather than genuine comprehension. Similar concerns have been raised in previous assessment research. For instance, studies on assessment quality emphasize that poorly designed distractors and heterogeneous option structures weaken the discriminative power of multiple-choice tests and reduce their effectiveness as measurement instruments. By redesigning the items to ensure grammatical homogeneity, logical option ordering, and plausible distractors, this study demonstrates how systematic item analysis can serve as a practical strategy for improving assessment quality.

Beyond structural improvements, this study also highlights the potential of digital platforms in supporting formative assessment. The reconstructed question bank was implemented using Quizizz, which enabled randomized item delivery, immediate feedback, and interactive engagement. Previous studies have shown that gamified assessment platforms such as Quizizz can enhance motivation, participation, and engagement in classroom learning. For example, Razali et al. (2020) found that gamification elements such as points and leaderboards significantly influence students' intrinsic and extrinsic motivation. Similarly, Melissa et al. (2022) reported that gamification features within Quizizz increase students' engagement and encourage active participation during quizzes. These findings suggest that digital assessment environments can transform traditional testing into more interactive learning experiences.

However, the literature also highlights potential challenges associated with gamified assessments. Melissa et al. (2022) noted that certain gamification features, such as leaderboards and speed-based scoring, may create anxiety or encourage rushed responses. This indicates that the design of gamified assessment should balance engagement with fairness and cognitive depth. In the present study, the emphasis was placed not only on gamification but also on improving the cognitive quality of assessment items. This approach reflects the view that technology alone cannot improve learning outcomes unless it is accompanied by strong pedagogical design. As Damayanti and Rahmadonna (2025) argue, digital tools are most effective when integrated within a structured instructional framework rather than used as standalone technological solutions.

The use of Quizizz in this study also aligns with research emphasizing the benefits of retrieval practice and formative quizzing. Experimental research by McDaniel et al. (2012) demonstrated that repeated quizzing significantly improves learning outcomes compared with passive study methods such as rereading. According to the testing effect theory, retrieving information through quizzes strengthens memory consolidation and improves long-term retention. By integrating the reconstructed question bank into a digital quizzing platform, this study leverages the cognitive benefits of retrieval practice while simultaneously enhancing engagement through interactive features.

From a broader perspective, the findings of this study contribute to the evolving landscape of digital assessment. Recent research highlights the increasing role of intelligent technologies in assessment design. For example, studies on AI-generated gamified quizzes suggest that automated assessment systems can increase participation rates, provide personalized feedback, and reduce instructor workload. These developments indicate a shift toward more adaptive and technology-driven assessment ecosystems. Within this context, the present study demonstrates how traditional question banks can be reconstructed into digitally supported formative assessment tools that promote both cognitive engagement and assessment quality.

These findings imply the need for systematic teacher training in item-writing skills, explicit integration of HOTS within curriculum planning, and strategic use of digital assessment tools not merely for engagement but for cognitive development. Despite these contributions, the study also has several limitations. First, the reconstruction process focused primarily on cognitive distribution and item construction rather than evaluating the empirical effectiveness of the reconstructed question bank on student learning outcomes. Future research could investigate the impact of the reconstructed question bank through experimental or quasi-experimental designs. Second, the study was limited to English question banks for junior high school, which may restrict the generalizability of the findings to other subjects or educational levels.

Overall, this study demonstrates that qualitative item analysis provides a systematic and practical approach for improving assessment quality. By integrating higher-order thinking skills and implementing the reconstructed items within a digital assessment platform, the study contributes to the development of more cognitively balanced and pedagogically meaningful assessment practices. These findings highlight the importance of aligning assessment design with contemporary educational goals that emphasize critical thinking, analytical reasoning, and active student engagement.

The shortcomings identified in the existing English question banks were addressed through a systematic reconstruction process guided by qualitative item analysis and established item-writing principles. The analysis of 60 multiple-choice items revealed several major weaknesses, including an imbalance in cognitive levels, structural flaws in item construction, and minor language inaccuracies. In terms of cognitive distribution, the majority of the existing items focused on lower-order thinking skills, particularly applying (C3), which accounted for 56.7% of the total items, while higher-order cognitive levels such as evaluating (C5) and creating (C6) were entirely absent. This imbalance indicates that the existing question banks primarily assessed procedural knowledge rather than analytical reasoning.

To address this issue, the reconstructed question bank redistributed the cognitive composition of the items to strengthen the integration of higher-order thinking skills. In the revised version, the proportion of analyzing (C4) items was

substantially increased, resulting in a new distribution consisting of four remembering items (C1), eight understanding items (C2), twelve applying items (C3), and sixteen analyzing items (C4). This redistribution promotes deeper cognitive engagement by requiring students to interpret information, identify relationships across texts, and draw logical conclusions.

In addition to cognitive restructuring, several construction-related weaknesses identified in the existing question banks were systematically corrected. A considerable number of items previously contained options that gave clues to the correct answer, heterogeneous grammatical structures, and implausible distractors. These weaknesses were addressed by redesigning the items to ensure that all options were homogeneous in grammatical structure, relatively equal in length, logically ordered, and supported by plausible distractors. The reconstructed items also clarified the stems to ensure that the central idea of the question was clearly stated in the stem rather than implied within the answer options.

The overall reconstruction process can be conceptualized as a three-stage model consisting of diagnostic analysis, reconstruction strategies, and enhanced assessment outcomes. As illustrated in Figure 1, the process begins with the qualitative analysis of the existing question bank to identify key shortcomings such as cognitive imbalance, poor item construction, and language-related issues. These findings then inform the reconstruction phase, which includes cognitive redistribution to increase higher-order thinking skills, improvement of item construction through better distractor design and clearer stems, and the integration of digital assessment tools. The final stage of the model represents the outcomes of the reconstruction process, namely a cognitively balanced question bank, improved item quality, a five-option multiple-choice format, and the implementation of a digital and interactive assessment environment.

Furthermore, the reconstructed question bank was implemented using the Quizizz platform to support digital formative assessment. The digital format enables randomized item presentation, automated scoring, and immediate feedback, which enhances assessment efficiency and student engagement. By integrating improved item design with digital delivery, the reconstructed question bank not only addresses the structural weaknesses identified in the original items but also supports more effective and interactive assessment practices in junior high school English learning.

Overall, the reconstruction process demonstrates that systematic qualitative item analysis can serve as a practical framework for improving both the cognitive quality and structural validity of assessment instruments. The integration of higher-order thinking skills and digital assessment tools further strengthens the role of the question bank as a meaningful formative assessment resource as shown on Table 4.

Table 4. Improvements Implemented in the Reconstructed Question Bank

Identified Shortcomings in Existing Question Banks	Evidence from Existing Items	Reconstruction Strategy	Improvement Outcome
Imbalanced cognitive distribution	Dominated by C3 (56.7%); no C5–C6 items	Redistributed cognitive levels	Increased HOTS representation with 16 C4 items
Options giving clues to the correct answer	22 items (36.7%)	Redesigned distractors and option structures	Eliminated clue-giving options
Heterogeneous grammatical structure in options	22 items (36.7%)	Standardized grammatical structure	Homogeneous options across all items
Unequal option length	17 items (28.3%)	Adjusted option length	Balanced option length
Lack of logical or numerical order	20 items (33.3%)	Organized options logically/alphabetically	Improved readability and clarity
Implausible distractors	14 items (23.3%)	Developed plausible distractors	Increased item discrimination
Unclear stems	8 items (13.3%)	Clarified main idea in stems	More focused question prompts
Limited distractor variation	Only 3 distractors per item	Increased options to five choices	Reduced guessing probability
Paper-based format	Traditional printed question bank	Implemented digital Quizizz platform	Randomization, instant feedback, and automated scoring

The findings of this study demonstrate that the shortcomings identified in the existing English question banks particularly cognitive imbalance and weaknesses in item construction can be systematically addressed through qualitative item analysis and the integration of higher-order thinking skills (HOTS) within a digitally supported assessment environment. The dominance of lower-order cognitive levels in the analyzed question banks reflects a pattern widely documented in previous assessment research.

For example, Abosalem (2016) found that teacher-developed tests tend to emphasize the lower levels of Bloom's taxonomy, particularly knowledge, comprehension, and application, while analytical and evaluative tasks are rarely included. Such patterns limit opportunities for students to develop critical thinking and complex reasoning abilities, suggesting that traditional assessment practices often fail to support higher-order cognitive development.

The reconstruction strategy implemented in this study directly responds to this challenge by redistributing the cognitive levels of assessment items and increasing the proportion of analytical tasks (C4). This approach aligns with the theoretical framework of Bloom's Revised Taxonomy, which emphasizes engaging students in higher cognitive processes such as analysis, evaluation, and creation. Previous empirical research also supports the importance of designing assessments that intentionally promote higher-order thinking.

For instance, Mahendra (2022) reported that integrating Quizizz-based formative assessment within project-based learning significantly improved students' higher-order thinking skills, including analytical reasoning and problem-solving abilities. Similarly, Rizki et al. (2022) found that the use of Quizizz as an assessment medium led to significantly higher post-test scores in higher-order thinking skills compared with other digital assessment platforms.

Beyond cognitive redistribution, the present study also addressed structural deficiencies in item construction. Poorly designed multiple-choice items such as those containing implausible distractors, heterogeneous option structures, or unclear stems can reduce item validity and allow students to guess answers through superficial cues rather than genuine comprehension. By redesigning the items to ensure plausible distractors, homogeneous grammatical structure, and clear stems, the reconstructed question bank improves both the validity and discriminative power of the assessment instrument. These findings are consistent with previous research on HOTS-based assessment development. For example, Dermawan et al. (2021) demonstrated that systematically designed HOTS-based instruments implemented through Quizizz achieved acceptable levels of reliability and validity when supported by rigorous development and validation procedures.

Another important dimension of the reconstruction process is the integration of digital formative assessment platforms. The implementation of the reconstructed question bank using Quizizz reflects the broader trend toward technology-enhanced assessment environments. Gamified digital tools have been widely recognized for their ability to increase engagement and motivation in learning activities. Razali et al. (2020) found that gamification elements such as points, rewards, and leaderboards significantly influence both intrinsic and extrinsic motivation among students. Similarly, Melissa et al. (2022) reported that gamification features in Quizizz increase engagement and competitiveness among students, although they may also introduce emotional pressure if not carefully managed.

These findings suggest that gamified assessment environments can enhance participation and motivation when designed with appropriate pedagogical considerations.

From a cognitive perspective, digital quizzing platforms also support retrieval-based learning processes. The experimental study conducted by McDaniel et al. (2012) provides strong empirical evidence for the “testing effect,” which suggests that repeated retrieval through quizzes strengthens long-term memory retention more effectively than passive study methods such as rereading. This theory highlights the role of formative quizzes as powerful learning tools rather than merely evaluative instruments. In the present study, the integration of Quizizz enables repeated retrieval practice through randomized question delivery and immediate feedback, thereby reinforcing both cognitive retention and formative learning.

Recent developments in educational technology research also highlight the growing role of intelligent systems in assessment design. Emerging studies on AI-generated gamified quizzes suggest that artificial intelligence can automate quiz generation while simultaneously enhancing engagement, participation, and learning outcomes. AI-supported gamified assessments can reduce instructor workload while providing scalable and adaptive assessment environments. Within this evolving landscape, the reconstructed question bank developed in this study represents a practical step toward more advanced digital assessment ecosystems by combining systematic item analysis, HOTS-oriented design, and gamified delivery platforms.

Despite these benefits, the literature also emphasizes that digital tools alone cannot guarantee meaningful learning outcomes. Damayanti and Rahmadonna (2025) argue that the effectiveness of gamified platforms such as Quizizz depends on their integration within well-designed instructional frameworks. Digital tools should therefore complement pedagogical strategies rather than replace them. In this regard, the reconstruction model proposed in this study highlights the importance of aligning cognitive design, item construction quality, and digital delivery mechanisms within a coherent formative assessment framework.

Thus, this study contributes to the growing body of research on digital formative assessment and HOTS-oriented evaluation practices. By addressing cognitive imbalance, improving item construction quality, and integrating gamified digital assessment platforms, the reconstructed question bank provides a systematic framework for enhancing both the cognitive rigor and pedagogical effectiveness of assessment practices in junior high school English learning. This study offers a practical reconstruction model that bridges the gap between assessment theory, item-writing practice, and digital formative assessment implementation.

5. Conclusion

The findings of this study have important implications for assessment practices in English language learning. The results indicate that many existing question banks still emphasize lower-order thinking skills, particularly at the remembering, understanding, and applying levels. Such an imbalance limits opportunities for students to develop higher-order thinking skills. Therefore, teachers and material developers need to redesign assessment items so that they better incorporate higher cognitive levels such as analysis and evaluation.

The reconstruction process applied in this study shows that qualitative item analysis can serve as a practical approach for improving both the cognitive level and structural quality of assessment items. In addition, the integration of digital assessment platforms such as Quizizz provides opportunities to enhance formative assessment practices. Digital quizzes enable randomized question delivery, immediate feedback, and interactive participation, which may increase student engagement and support more effective classroom assessment.

However, this study has several limitations. The research primarily focused on the qualitative analysis and reconstruction of question bank items without empirically testing the effectiveness of the reconstructed items in real classroom settings. As a result, the study demonstrates improvements in assessment design but does not measure their direct impact on students' learning outcomes or higher-order thinking development. Furthermore, the analysis was limited to question banks from two sources, which may not fully represent the diversity of English assessment practices in junior high schools. These limitations suggest that the findings should be interpreted within the specific context of the analyzed materials.

Future research may extend this study by examining the effectiveness of the reconstructed question bank through classroom implementation and experimental or quasi-experimental research designs. Such studies could investigate whether the improved item design contributes to measurable improvements in students' higher-order thinking skills and academic performance. Further research may also expand the scope of analysis by including question banks from different subjects, grade levels, and educational contexts to provide a broader understanding of assessment practices. In addition, future studies could explore the integration of digital assessment tools and emerging technologies to support more innovative and effective formative assessment environments.

6. References

- Al-reshaid, N., & Alhojailan, A. I. (2025). Stylistic Variations in Thematic Structure Across Academic Genres : A Case Study of EFL Graduate Students. 15.
- Alqasham, F. H., Al-Ahdal, A. A. M., & Babekir, A. H. S. (2021). Coherence and Cohesion in Saudi EFL Learners' Essay Writing: A Study at a Tertiary-level Institution. *Asian EFL Journal*, 28(1), 8–25. https://www.researchgate.net/publication/349256060_Coherence_and_Cohesion_in_Saudi_EFL_Learners'_Essay_Writing_A_Study_at_a_Tertiary-level_Institution
- Aminova, B. D. (2022). Importance Of Coherence and Cohesion in Writing. *Eurasian Research Bulletin*, 4, 83–89.
- Amoah, S., & Yeboah, J. (2021). The speaking difficulties of Chinese EFL learners and their motivation towards speaking the English language. *Journal of Language and Linguistic Studies*, 17(1), 56–69. <https://doi.org/10.52462/jlls.4>
- Asep Saepuloh, Ningsih, D. N., & Halimah. (2023). Theme-Rheme And Thematic Progression Analysis Of Joe Biden's Speech On Israel-Hamas And Russia-Ukraine Wars. *International Journal of Educational and Life Sciences*, 1(4), 217–228. <https://doi.org/10.59890/ijels.v1i4.991>
- Brian Paltridge. (2022). *Discourse An Analysis : An Introduction* (K. Hyland (ed.); 3rd ed.).
- Canestra, N., & Putra, A. (2023). STUDENTS ' ANXIETY IN SPEAKING ENGLISH IN. 10(2), 108–120. <https://doi.org/10.22219/celtic.v10i2>.
- Dashela, T. (2021). The Analysis of Theme and Rheme in Short Story of Sleeping Beauty with a Systemic Functional Approach. *SALEE: Study of Applied Linguistics and English Education*, 2(1), 11–28. <https://doi.org/10.35961/salee.v2i01.201>
- Daymiel, R. J., Cantina, J. M., Alipoyo, V. R. I., Comecilla, M. O., Patay, A. S., & Recamara, J. T. (2022). Anxiety in Second Language in Relation to Students' Speaking Performance. *Sprin Journal of Arts, Humanities and Social Sciences*, 01(August), 396–410. <https://doi.org/10.55559/sjahss.v1i08.41>
- Eggin, S. (2004). *An Introduction to Systemic Functional Linguistics*. Continuum International Publishing Group. <https://doi.org/https://doi.org/10.1111/j.1360-6441.2006.0327j.x>
- Emilia, E., & Hamied, F. A. (2015). SYSTEMIC FUNCTIONAL LINGUISTIC GENRE PEDAGOGY (SFL GP) IN A TERTIARY EFL WRITING CONTEXT IN INDONESIA. *TEFLIN Journal*, 26(2), 155–182. <https://doi.org/https://doi.org/10.15639/teflinjournal.v26i2/155-182>
- Gerot, L., & Wignell, P. (1994). *Making Sense of Functional Grammar. An introductory workbook*. Antipodean Educational Enterprises.
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Halliday's Introduction to Functional Grammar* (4th ed.). Routledge.

- Hameed, Y. S., & Challob, A. I. (2022). Synonymy as a Cohesive Device in Students' English Essay Writing. *Al-Adab Journal*, 140, 91–110. <https://doi.org/10.31973/aj.v2i140.3632>
- Happy gusnar, N. S., Hartono, R., & Suwandi, S. (2021). Cohesion and Coherence in Final Project Abstracts Written by Non-English Department Undergraduate Students. *English Education Journal*, 11(4), 516–527. <https://doi.org/10.15294/eej.v11i1.47453>
- Hendrawan, W., Sujatna, E. T. S., Krisnawati, E., & Darmayanti, N. (2021). Thematic Progression in Uncovering Paragraph Coherence: Some Evidences in Indonesian Secretarial Academy Context. *TESOL International Journal*, 16(1), 179–209.
- Himphinit, M., & Astia, I. (2023). Systemic Functional Linguistics : Analyzing the Theme System and Thematic Progression in Indonesian University Students ' English (Vol. 1). Atlantis Press SARL. <https://doi.org/10.2991/978-2-38476-022-0>
- Islami, K. H., Saleh, M., & Linggar Bharati, D. A. (2022). The Use of Cohesive Devices in Descriptive Text by English Training Participants at PST". *English Education Journal*, 12(1), 96–103. <https://doi.org/10.15294/eej.v12i1.49052>
- Jiang, Y., Chen, Y., Lu, J., & Wang, Y. (2021). The Effect of the Online and Offline Blended Teaching Mode on English as a Foreign Language Learners' Listening Performance in a Chinese Context. *Frontiers in Psychology*, 12(November). <https://doi.org/10.3389/fpsyg.2021.742742>
- Knapp, B. (2024). Skill in Sport: The Attainment of Proficiency. *Skill in Sport: The Attainment of Proficiency*, 1–204. <https://doi.org/10.4324/9781003484110/SKILL-SPORT-KNAPP/RIGHTS-AND-PERMISSIONS>
- Larasati, K. D., I Gusti Ayu Vina Widiadnya Putri, & Putu Nur Ayomi. (2022). Analysis on Thematic Progression Patterns of Boris Johnson's Speech. *ELYSIAN JOURNAL: English Literature, Linguistics and Translation Studies*, 2(4), 100–109. <https://doi.org/10.36733/elysian.v2i4.4400>
- Meng, J. (2022). A Comparative Study on the Master Position Advancement Mode of English Abstracts for Linguistics Dissertations. *Journal of Hefei College (General Edition)*, 39(03), 60–65.
- Ngongo, M., & Ndandara, A. (2024). Exploring the Realization of Cohesion As a Resource of Text Texture on Undergraduate Students' Academic Writing. *English Review: Journal of English Education*, 12(1), 343–352. <https://doi.org/10.25134/erjee.v12i1.9241>
- Purwati, D., Faruq Ubaidillah, M., & Restall, G. C. (2023). "Sorry, I Can't Speak": English Teachers' Challenges of Teaching EFL Speaking in an Indonesian Vocational High School Sector. *Mextesol Journal*, 47(1), 0–2. <https://doi.org/10.61871/mj.v47n1-1>

- Putra, P. P., Fatmawati, U., Bengkulu, S., & Astari, Y. (2022). Analysis of Cohesion and Coherence in the Background of Thesis Written by English Education Study Program Students of IAIN Bengkulu in Period September 2020. *Linguistics and Literature*, 5(1), 79–90.
- Rizki, M., Hidayat, D. N., Husna, N., & Alex. (2022). DISCOURSE ANALYSIS OF GRAMMATICAL COHESION DEVICES IN STUDENT EXPLANATION. *Journal of English Education and Development*, 6(1), 1–12. <https://doi.org/https://doi.org/10.31605/eduvelop.v6i1.1545>
- Rodliyah, R. S., & Liani, A. E. (2022). SFL analysis: An investigation of students' use of cohesive devices in exposition text. *Indonesian Journal of Applied Linguistics*, 12(1), 243–254. <https://doi.org/10.17509/ijal.v12i1.46596>
- Rusdin, R., & Purwati, D. (2023). Speaking Up: A Comprehensive Investigation of EFL Secondary Students' Speaking Skill Problems in Indonesia's Islamic School Context. *Elsya : Journal of English Language Studies*, 5(2), 236–249. <https://doi.org/10.31849/elsya.v5i2.11911>
- Siregar, J., Nurlela, N., & Zein, T. T. (2023). An Analysis of Cohesive Devices in EFL Students' Essay Writing. *Indonesian Journal of EFL and Linguistics*, 39–53. <https://doi.org/10.21462/IJEFL.V8I1.577>
- Sulastri, S. (2022). How A Student Develops Paragraphs : A Thematic Progression Analysis of Student Works on Short Story. *Journal Of Education And Teaching Learning (JETL)*, 4(2), 105–114. <https://doi.org/10.51178/jetl.v4i2.558>
- Trinh, N. B., & Pham, D. T. T. (2021). Challenges in speaking classrooms among non-English majors. *Vietnam Journal of Education*, 5(2), 26–31. <https://doi.org/10.52296/vje.2021.96>
- Turner, J. E., Li, B., & Wei, M. (2021). Exploring effects of culture on students' achievement motives and goals, self-efficacy, and willingness for public performances: The case of Chinese students' speaking English in class. *Learning and Individual Differences*, 85(December 2020), 101943. <https://doi.org/10.1016/j.lindif.2020.101943>
- Utama, A. R., Prabowo, R. K., Lusiana Indra Sari, & Kurniawan, A. (2026). A Textual Metafunction Analysis of Vocational High School Students' Speeches. 97–124.
- Utama, A. R., Sri Wuli Fitriati, & Zulfa Sakhiyya. (2024). Unveiling Cohesion in Senior High School Students' Speeches: A Systemic Functional Linguistics Approach. *SALEE: Study of Applied Linguistics and English Education*, 5(2), 603–622. <https://doi.org/10.35961/salee.v5i2.1519>
- Wakerkwa, D. A. P. (2024). GRAMMATICAL COHESIVE FEATURES IN INDONESIAN EFL STUDENTS' ARGUMENTATIVE ESSAYS. *ACCELERATION: Multidisciplinary Research Journal*, 02(01), 56–63. <https://doi.org/https://doi.org/10.70210/amrj.v2i01.63>
- Wang, T. (2021). Discourse Analysis of Lord of the Flies : A Systemic Functional

- Approach. *Advances in Literary Study*, 9(3), 114–126.
<https://doi.org/10.4236/als.2021.93013>
- Wijayanti, H., Suwandi, S., & Fitriati, S. W. (2023). Assessing the Students' Use of Cohesive Devices in Writing Hortatory Exposition Text. *English Education Journal*, 13(3), 363–369. <https://doi.org/10.15294/eej.v13i3.72805>
- Zahra, G. M., Emilia, E., & Nurlaelawati, I. (2021). An Analysis of Cohesion and Coherence of Descriptive Texts Written by Junior High School Students. *Proceedings of the Thirteenth Conference on Applied Linguistics (CONAPLIN 2020)*, 546(Conaplin 2020), 195–202. <https://doi.org/10.2991/assehr.k.210427.030>
- Zamroni, A., Wuli Fitriati, S., & Widhiyanto, W. (2023). The Realization of Thematic Progression in Background of Study Section Written by Undergraduate Students at Tidar University. *English Education Journal*, 13(2), 162–169. <https://doi.org/10.15294/eej.v13i2.71502>