Iranian Journal of Applied Linguistics (IJAL)
Vol. 27, No. 2, September 2024, 25-47

Aligning Learning Objectives in *Select Readings* with University Achievement Tests Using Bloom 's Revised Taxonomy

Sahar Zahed Alavi*
University of Bojnord, Bojnord, Iran.
Mohammad Reza Ghorbani
University of Bojnord, Bojnord, Iran.

Abstract

In educational settings, a common concern is whether instructional materials and assessments are cognitively aligned to foster deeper learning. This study investigates the alignment between the learning objectives presented in the *Select Readings* textbook and those evaluated in university-standardized achievement tests at Shiraz University, using Bloom's Revised Taxonomy (2001) as the analytical lens. Drawing on Razmjoo and Kazempurfard's (2012) coding methodology, both the textbook and two different test forms from different semesters were systematically analyzed. The results revealed a clear dominance of lower-order thinking skills, particularly factual knowledge within the Remembering and Understanding domains. Notably, the achievement tests emphasized Understanding to a greater extent than the textbook. Although the textbook incorporated some higher-order cognitive domains (Analyzing, Evaluating, Creating), these appeared infrequently in the tests. The findings point to a misalignment between instructional intent and assessment focus, raising concerns about the depth and validity of the evaluation tools. Contextual variables such as instructional intent, learner proficiency, and curricular constraints may also contribute to this imbalance, warranting further examination. The study underscores the need for assessment designers to adopt more cognitively diverse items that reflect the full spectrum of learning objectives embedded in teaching materials.

Keywords: Learning Objectives, Textbook Analysis, Assessment, Cognitive Levels, Bloom's Revised Taxonomy

* *Corresponding author: University of Bojnord, Iran.* *Email address:* s.zahedalavi@ub.ac.ir

[DOR: 20.1001.1.17351634.2024.27.2.2.6]

1. Introduction

In language education, textbooks play a central role in shaping instructional content, especially in settings where they serve as the primary source of linguistic input (Para, 2024; Razmjoo, 2010; Shahmohammadi, 2018). As Tomlinson (2011) emphasizes, textbooks are not merely collections of exercises - they function as structured pedagogical tools intended to meet learners' educational needs, guiding instruction while promoting consistency and quality (McGrath, 2002; Kostas, 2023; Richards & Renandya, 2002).

Evaluating these materials is essential to ensure their alignment with broader curriculum goals. Poor textbook selection can lead to diminished learner engagement, ineffective teaching, and wasted resources (Guo & Yao, 2021; McDonough & Shaw, 2003; Mukundan, 2007). Equally vital is the role of assessment, particularly summative evaluations like standardized achievement tests, which aim to measure learning outcomes (Uskokovic & Schirm, 2024; Bachman & Palmer, 2010). Ideally, assessments should encompass a range of cognitive demands - from foundational recall to advanced problem-solving - to accurately reflect student understanding (Le & Chong, 2024; Freahat & Smadi, 2014).

A growing body of research has examined textbooks and assessments through the lens of Bloom's Taxonomy, revealing that educational materials and tests often emphasize lower-order cognitive skills at the expense of higher-order thinking skills (e.g., Divsar & Jafarigohar, 2020; Freahat & Smadi, 2014; Chavda et al., 2023; Le & Chong, 2024). However, most previous studies have focused either on textbook content or assessment practices independently, with relatively limited attention given to the degree of alignment between learning objectives embedded in instructional materials and those assessed in achievement examinations. Furthermore, studies investigating such alignment in English as a Foreign Language (EFL) contexts, particularly in Iranian higher education settings, remain scarce. This gap is important because misalignment between instructional objectives and assessment practices may result in students being evaluated on cognitive skills that are insufficiently emphasized during instruction, thereby weakening the validity of assessment outcomes and the effectiveness of teaching and learning processes.

[DOR: 20.1001.1.17351634.2024.27.2.2.6]

This study is grounded in Bloom's Revised Taxonomy (2001), a framework that expands on the original taxonomy by introducing a two-dimensional model: the cognitive domain and the knowledge domain. The original taxonomy identified six sequential cognitive processes: Knowledge, Comprehension, Application, Analysis, Synthesis, and Evaluation. The revised version reorganizes these into six updated categories that reflect contemporary learning theory:

- Remembering: retrieving and recognizing learned information;
- Understanding: interpreting, summarizing, and comparing concepts;
- Applying: using knowledge in familiar or novel contexts;
- Analyzing: breaking information into parts and uncovering relationships;
- Evaluating: making judgments using criteria and standards;
- Creating: combining elements to form coherent or original products.

Alongside these cognitive processes, the revised taxonomy incorporates four knowledge types (Krathwohl, 2002):

- Factual: essential details and terminology;
- Conceptual: categories, principles, and theories;
- Procedural: techniques and methods for task execution;
- Metacognitive: awareness of learning strategies and self-regulation.

This dual-dimensional framework offers a nuanced lens for analyzing instructional content and assessments (Chavda et al., 2023). It allows educators to evaluate whether educational materials foster a balanced distribution of thinking skills - and whether assessments accurately measure these objectives. Such alignment has implications for student performance, as disproportionate emphasis on lower-order skills may hinder learners' capacity for deeper cognitive engagement.

The need for such investigation is particularly relevant in the Iranian university context, where achievement tests often play a decisive role in determining students' academic success and are frequently developed by course instructors based on prescribed textbooks. However, little empirical evidence exists regarding whether the cognitive demands embedded in the textbook are adequately reflected in university achievement tests. This lack of evidence represents a local pedagogical and assessment concern because potential discrepancies between instructional content and testing practices may influence both teaching priorities and students' opportunities to develop higher-order thinking skills. Therefore, examining the

[DOR: 20.1001.1.17351634.2024.27.2.2.6]

alignment between textbook learning objectives and achievement tests can provide valuable insights for improving curriculum implementation, assessment validity, and instructional quality within this context.

Therefore, this study examines the cognitive levels represented in the *Select Readings* textbook (pre-intermediate level) and the corresponding university achievement tests administered at Shiraz University. The research seeks to answer the following questions:

1. How are Bloom's Revised Taxonomy learning objectives reflected in the *Select Readings* textbook and the university achievement tests?
2. Which levels of Bloom's Taxonomy are most frequently emphasized?
3. Do the assessments adequately reflect the higher-order thinking skills present in the textbook?

2. Review of related literature

The application of Bloom's Original and Revised Taxonomies has become a foundational tool for analyzing the cognitive demands of educational materials and assessments. Numerous studies (e.g., Assaly & Smadi, 2015; Chavda et al., 2023; Mizbani & Chalack, 2017; Razmjoo & Kazempur, 2012; Sahragard & Alavi, 2016; Stevani & Tarigan, 2022; Veeravagu et al., 2010) have explored how well textbooks and exams engage learners across cognitive levels, particularly the balance between lower-order and higher-order thinking skills. These findings consistently reveal a predominance of lower-order cognitive activities - such as Remembering and Understanding - with minimal integration of higher-order processes like Evaluating and Creating.

2.1. Textbook evaluations based on Bloom's Taxonomy

Studies investigating textbook content frequently highlight an imbalance across cognitive domains. For example, Stevani and Tarigan (2022) analyzed two popular textbooks (Active Reading and English in Mind) and found that Understanding tasks made up 26% of all activities, while Creating tasks were nearly absent. This suggests that learners were rarely challenged beyond basic comprehension.

Similarly, Sahragard and Alavi (2016) compared textbooks written by native versus non-native English authors. Although lower-order thinking skills dominated both types, the non-native-authored

[DOR: 20.1001.1.17351634.2024.27.2.2.6]

textbooks featured a slightly higher incidence of higher-order tasks, such as Evaluating and Creating, indicating a more cognitively diverse instructional approach.

Using Ganbari's (2013) framework, Askaripour (2014) reviewed the Top-Notch series, identifying significant gaps in critical thinking. While Remembering, Understanding, and Applying were common, the cognitive level of Evaluating was entirely absent. This lack of critical judgment tasks led researchers to question the textbooks' ability to foster deep learning.

Razmjoo and Kazempur (2012) conducted a multi-edition analysis of the Interchange series, showing excessive reliance on lower-order thinking, particularly factual recall. They also observed inconsistencies across editions, with little attention paid to Metacognitive Knowledge, further limiting the textbooks' developmental potential.

In a high-school context, Mizbani and Chalack (2017) evaluated the Prospect 3 textbook's reading and writing activities. Their results emphasized a dominance of factual and procedural knowledge within Remembering and Understanding categories. Advanced cognitive skills such as Analyzing or Creating were not represented, prompting calls for curriculum supplementation.

Further supporting these findings, Assaly and Smadi (2015) reviewed reading comprehension questions in the Master Class textbook for tenth-grade learners. While Understanding accounted for over half of the tasks (52%), Remembering and Applying were notably underrepresented. Although 40% of questions involved higher-order skills, the researchers concluded that their distribution was uneven and ineffective for promoting consistent critical engagement.

2.2. Cognitive levels in language assessments

The same imbalance has been observed in standardized testing. Koksal and Ulum (2018) examined university-level General English exams, concluding that Remembering and Understanding were the only cognitive processes consistently represented. No assessment items targeted Analyzing, Evaluating, or Creating, suggesting a narrow view of learner competence.

Razmjoo and Madani (2013) assessed Iran's national Konkur university entrance exams, which are influential in shaping instruction. The researchers found near-exclusive focus on lower-order skills,

[DOR: 20.1001.1.17351634.2024.27.2.2.6]

with Creating being completely absent - a significant limitation in assessing critical thinking and problem-solving abilities.

In the high school context, Kamlasi and Sahan (2018) evaluated final exams and noted an overrepresentation of Remembering (44%) and Applying (42%), while Understanding constituted only 4% and Analyzing appeared in just 10% of items. They recommended test development practices that better balance cognitive demands.

Baghaei et al. (2020) analyzed IELTS and TOEFL exams for cognitive diversity. IELTS listening tasks centered on Understanding and Remembering, whereas TOEFL showed slightly greater variety, including Analyzing and Creating in the listening section. However, the reading components in both remained mostly confined to lower-order cognitive functions, raising concerns about whether high-stakes tests encourage deep learning.

2.3. Research gap and rationale

While previous studies have provided extensive analyses of either textbook content or test items using Bloom's Revised Taxonomy, few have directly compared the learning objectives of a textbook with the cognitive distribution of assessments based explicitly on that textbook. This is a crucial oversight, as such comparisons offer insights into curriculum coherence, instructional quality, and assessment validity.

Emerging research on curriculum analytics, such as Jovanović et al. (2025), has begun to address these concerns through semi-automated mapping of learning objectives to assessment tasks. Their findings highlight the risks of instructional misalignment and suggest the need for frameworks that systematically evaluate the cognitive congruence between course materials and assessments. Similarly, Li and Wang (2024) emphasize that misaligned textbooks and assessments negatively influence student performance, advocating for integrated design strategies.

Despite these developments, whether university-level achievement tests based on a specific textbook appropriately reflect the cognitive goals embedded in that textbook remains largely unexplored. This lack of alignment risks evaluating students on skills they were not adequately prepared to master, undermining the integrity of assessment practices.

[DOR: 20.1001.1.17351634.2024.27.2.2.6]

Addressing this gap is vital for promoting reflective curriculum design. Ensuring textbook-test alignment means that assessments can validly measure learners' progress across a range of cognitive processes - from foundational recall to critical synthesis - ultimately enhancing educational outcomes in higher education.

3. Methodology

3.1. Research design

This study employed a qualitative content analysis design, guided by Bloom's Revised Taxonomy (2001), to systematically investigate the cognitive and knowledge-related features of instructional materials and their corresponding assessments. Content analysis enables researchers to interpret textual data through pre-established theoretical categories (Ary et al., 2019), allowing for a deeper understanding of the educational objectives embedded within learning tasks.

Although qualitative in its foundation, the study incorporated quantitative techniques - such as frequency counts and chi-square testing - rendering it a convergent mixed-methods design. This approach provided both interpretive richness and empirical robustness in assessing cognitive demands across instructional sources.

The study utilized purposive sampling to select instructional materials from a General English course at Shiraz University. The sample consisted of:

Select Readings (Pre-Intermediate Level), a widely used general English textbook in the university's curriculum.

Two official summative achievement tests, created by the university's English department, designed to assess vocabulary and reading comprehension based on the textbook content.

To ensure representative coverage while maintaining data manageability, seven chapters were randomly selected from the textbook. Randomization within a purposive framework helped mitigate researcher bias and ensured a diverse range of pedagogical activities were captured. These chapters featured various task types, including: pre-reading prompts, reading passages, comprehension questions, vocabulary exercises, grammar tasks, and post-reading discussions.

[DOR: 20.1001.1.17351634.2024.27.2.2.6]

The selection of seven chapters was based on methodological considerations commonly adopted in qualitative content analysis, where the objective is to obtain sufficient representativeness of content categories rather than statistical generalization (Kuckartz, 2025). Previous textbook evaluation studies employing Bloom's Taxonomy have likewise analyzed selected units or chapters rather than entire textbooks due to the intensive nature of item-level coding (e.g., Razmjoo & Kazempur, 2012; Mahmood et al., 2020). The randomly selected chapters represented approximately 50% of the textbook and included all major instructional components recurring throughout the book. Preliminary examination further indicated that the selected chapters contained a broad distribution of task types and learning objectives comparable to those found in the remaining chapters. Therefore, the sample was considered sufficiently representative for identifying the textbook's overall cognitive profile while maintaining analytical feasibility.

3.2. Instrument and coding scheme

The analysis employed the coding framework introduced by Razmjoo and Kazempur (2012), which operationalizes Bloom's Revised Taxonomy across two core dimensions:

Cognitive Process Dimension:

A – Remembering

B – Understanding

C – Applying

D – Analyzing

E – Evaluating

F – Creating

Knowledge Domain Dimension:

1 – Factual Knowledge

2 – Conceptual Knowledge

3 – Procedural Knowledge

[DOR: 20.1001.1.17351634.2024.27.2.2.6]

4 – Metacognitive Knowledge

Each instructional activity and test item was assigned a composite code (e.g., B1 = Understanding Factual Knowledge) based on the nature of the task and its intended learning objective.

The coding process followed a deductive content analysis approach (Elo & Kyngäs, 2008), as all categories were established a priori from Bloom's Revised Taxonomy and the coding framework proposed by Razmjoo and Kazempur (2012). Rather than generating categories inductively from the data, the analysis sought to determine the extent to which textbook tasks and test items reflected predetermined cognitive and knowledge dimensions.

The unit of analysis consisted of individual textbook activities and individual test items. Coding was conducted primarily at the manifest level, focusing on explicitly stated task requirements, instructional prompts, and observable learning outcomes .

In addition, the coding scheme primarily involved descriptive coding, whereby each activity or item was categorized according to its dominant cognitive process and knowledge type. At the same time, interpretive judgment was required to determine the underlying cognitive demand represented by a task. For example, an item requiring students to compare ideas was coded as Analyzing, whereas an item requiring justification of a position was coded as Evaluating. To maintain consistency, detailed coding guidelines and exemplar cases from the original framework were consulted throughout the analysis.

The coding process was conducted manually by the primary researcher, using a standardized rubric. Ambiguous items were noted for peer consultation, and all coded entries were documented in Excel spreadsheets for systematic analysis. The full coding scheme is provided in the study's Appendix.

3.3. Instrument reliability and validity

To ensure coding reliability and credibility, both intra-coder and inter-coder reliability procedures were implemented:

Intra-Coder Reliability:

The primary researcher re-coded a sample of three chapters and 20 test items after a one-month interval. A consistency rate of 95.4% was achieved, indicating strong internal reliability.

[DOR: 20.1001.1.17351634.2024.27.2.2.6]

Inter-Coder Reliability:

An assistant professor of TEFL at Tehran University, trained in Bloom's Revised Taxonomy, independently coded the same sample using the same rubric. The resulting inter-coder agreement was 93.5%, calculated via SPSS using standard statistical methods.

Validity of the coding instrument was enhanced through expert review and theoretical triangulation. Two applied linguistics scholars evaluated the coding rubric to ensure interpretive alignment with educational objectives. Furthermore, the use of Bloom's Revised Taxonomy, a widely validated theoretical framework, strengthened the construct validity by linking item classification to well-established cognitive categories.

3.4. Data collection and data analysis procedures

The dataset for this study was obtained from two primary sources: printed copies of the instructional textbook and two different test forms from different semesters. Textbook chapters relevant to the instructional aims of the curriculum were manually reviewed, and all embedded tasks were extracted and labeled according to their pedagogical function - such as vocabulary reinforcement, reading comprehension, or grammar practice. Concurrently, each test item was examined for cognitive rigor and targeted knowledge domain.

A systematic coding process was employed using a standardized rubric, allowing for consistent classification across both materials. This rubric captured dimensions of cognitive demand (e.g., recall, application, evaluation) and knowledge type (e.g., factual, procedural, conceptual). All coded data were compiled and entered into Excel spreadsheets, preparing the groundwork for subsequent quantitative analyses.

Although the initial data collection was qualitative in nature, the study integrated quantitative techniques to identify and interpret instructional trends. Frequency counts of cognitive and knowledge codes were aggregated for both textbook and test datasets. These counts were then used for comparative analysis to evaluate the presence and balance of lower-order versus higher-order thinking skills across the two sources.

[DOR: 20.1001.1.17351634.2024.27.2.2.6]

To assess whether differences in cognitive representation between textbook tasks and test items were statistically meaningful, chi-square tests were conducted via SPSS. This inferential analysis determined the likelihood that the observed discrepancies were attributable to instructional design rather than random variance, thereby offering insights into potential systemic pedagogical patterns.

Ethical rigor was maintained throughout the study. All data sources comprised publicly accessible educational materials, with no involvement of personal or sensitive learner data. The second coder engaged voluntarily and was fully briefed on the study's objectives and protocols, ensuring transparency and ethical compliance in coding and reporting procedures.

4. Results

The analysis of *Select Readings* and university achievement tests revealed distinct patterns in the distribution of learning objectives based on Bloom's Revised Taxonomy (2001). The findings are presented in three stages: descriptive frequency analysis, inferential chi-square testing, and cognitive skill categorization.

4.1. Cognitive distribution in *Select Readings* and achievement tests

Table 1 displays the frequencies and percentages of learning objectives identified in both instructional materials. In Select Readings, the most frequent objective was Understanding factual knowledge (37%), closely followed by Remembering factual knowledge (36.22%). Higher-order cognitive processes - such as Analyzing, Creating, Applying, and Evaluating - were present but notably underrepresented. No metacognitive objectives were observed.

Achievement tests showed an even stronger bias toward lower-order thinking, with Understanding factual knowledge comprising 77% of all items, and Remembering factual knowledge accounting for 15%. Higher-order objectives such as Evaluating (5%) and Analyzing (3%) appeared minimally, while Applying and Creating were entirely absent.

[DOR: 20.1001.1.17351634.2024.27.2.2.6]

Table 1

Frequencies and percentages of learning objectives in Select Readings and achievement tests

Learning Objective	Select Readings	Achievement Tests
Remembering		
A1 (Factual)	46 (36.22%)	15 (15%)
A2 (Conceptual)	1 (0.79%)	0 (0%)
A3 (Procedural)	0 (0%)	0 (0%)
A4 (Metacognitive)	0 (0%)	0 (0%)
Understanding		
B1 (Factual)	47 (37%)	77 (77%)
B2–B4	0 (0%)	0 (0%)
Applying		
C1 (Factual)	4 (3.15%)	0 (0%)
C2 (Conceptual)	7 (5.51%)	0 (0%)
C3–C4	0 (0%)	0 (0%)
Analyzing	10 (7.87%)	3 (3%)
Evaluating	5 (3.94%)	5 (5%)
Creating	7 (5.51%)	0 (0%)
Metacognitive (D4, E4, F4)	0 (0%)	0 (0%)

4.2. Chi-Square analysis of learning objectives

To determine whether the observed differences in learning objective distributions were statistically significant, Chi-square tests were conducted. Results indicated significant variance in both datasets, suggesting that the distribution of cognitive levels was not random.

Table 2

Chi-Square test results for learning objectives

Source	χ^2	df	p
Select Readings	130.02	6	< .001
Achievement Tests	147.52	3	< .001

[DOR: 20.1001.1.17351634.2024.27.2.2.6]

4.3. Lower-order vs. higher-order thinking skills

Learning objectives were further grouped into lower-order skills (Remembering, Understanding, Applying) and higher-order skills (Analyzing, Evaluating, Creating). Table 3 shows that both materials heavily favored lower-order cognitive engagement.

Table 3

Distribution of lower and higher-order cognitive skills

Source	Lower-Order (%)	Higher-Order (%)
Select Readings	82.68% (105 instances)	17.32% (22 instances)
Achievement Tests	92% (92 instances)	8% (8 instances)

4.4. Chi-Square test for cognitive skill distribution

A second Chi-square analysis confirmed that the imbalance between lower- and higher-order skills was statistically significant.

Table 4

Chi-Square test results for cognitive skill distribution

Source	χ^2	df	p
Select Readings	54.24	1	< .001
Achievement Tests	70.56	1	< .001

4.5. Summary of findings

The results demonstrate a pronounced emphasis on recall-based learning in both instructional and assessment materials. While Understanding and Remembering dominate, higher-order cognitive processes - such as Analyzing, Evaluating, and Creating - are significantly underrepresented. The statistical analyses confirm that these disparities are not due to chance, underscoring the need for curricular realignment to promote deeper cognitive engagement and critical thinking.

[DOR: 20.1001.1.17351634.2024.27.2.2.6]

5. Discussion

The findings of this study reveal a notable cognitive misalignment between the learning objectives embedded in *Select Readings* and those assessed in its corresponding achievement tests. As illustrated in Table 1, the textbook spans the full spectrum of Bloom's Revised Taxonomy - Remembering, Understanding, Applying, Analyzing, Evaluating, and Creating - whereas the achievement tests assess only four of these levels, omitting Applying and Creating entirely. This discrepancy undermines the principle of constructive alignment (Biggs, 2003), which posits that instructional materials, learning activities, and assessments should be coherently aligned to support intended learning outcomes.

Such an imbalance contradicts established principles of assessment design (Koksal & Ulum, 2018; Kosasih et al., 2022; Kowalski et al., 2024; Omar et al., 2012; Salas et al., 2024), which advocate for assessments that reflect a hierarchical progression of cognitive complexity. However, this divergence may be partially attributed to the nature of instructional tasks. Textbook activities often involve productive skills - such as speaking, writing, and collaborative discussion - that naturally elicit Applying and Creating (Astrid et al., 2022; Littlejohn, 2011). In contrast, achievement tests, designed for summative evaluation, tend to favor standardized formats that prioritize reliability and ease of scoring, often at the expense of cognitive depth (Shepard, 2000).

5.1. Impact of test format on learning objectives

The exclusive reliance on multiple-choice items in the achievement tests further explains the absence of Applying and Creating. While multiple-choice formats offer logistical advantages - objectivity, efficiency, and scalability - they inherently constrain students to recognition-based responses, limiting opportunities for idea generation, synthesis, and problem-solving (Palmer & Devitt, 2006; Quana et al., 2018; Veeravagu, 2010). In contrast, the textbook's inclusion of open-ended tasks fosters productive engagement, allowing learners to demonstrate procedural fluency and creative reasoning.

Moreover, the sequencing of textbook activities - from pre-reading to post-reading - may influence cognitive engagement. Pre-reading tasks often activate prior knowledge and encourage prediction or hypothesis formation, aligning with higher-order skills. Post-reading tasks, however, tend to reinforce comprehension and recall (Van Loi & Thanh, 2022). Since achievement tests predominantly assess post-reading comprehension, they inadvertently reinforce lower-order cognitive processing.

[DOR: 20.1001.1.17351634.2024.27.2.2.6]

5.2. Imbalance between understanding and remembering

A further asymmetry is observed in the distribution of lower-order cognitive skills. While the textbook maintains a near-equitable balance between Remembering (36.22%) and Understanding (37%), the achievement tests disproportionately emphasize Understanding (77%) over Remembering (15%). This suggests that test designers may be targeting conceptual comprehension rather than rote memorization, possibly reflecting an assumption that students, having completed the textbook, should demonstrate deeper understanding. However, this shift within the lower-order spectrum may not compensate for the absence of higher-order cognitive demands, particularly in contexts where critical thinking and transferability are essential.

5.3. Neglect of procedural and metacognitive knowledge

Beyond cognitive levels, the study highlights a limited representation of procedural and metacognitive knowledge across both instructional and assessment materials. While factual and conceptual domains - such as vocabulary, grammar, and text structure - are adequately addressed, strategic knowledge (e.g., how to approach a task, monitor comprehension, or evaluate one's own performance) remains underdeveloped. This omission is concerning, given that procedural and metacognitive skills are pivotal for autonomous learning and self-regulation (Krathwohl, 2002; Schraw & Moshman, 1995). The findings echo prior research (Sahragard & Alavi, 2016; Razmjoo & Kazempourfard, 2012), which cautions that EFL learners may struggle to transfer knowledge or adapt strategies in novel contexts without explicit scaffolding.

5.4. Lower-order vs. higher-order cognitive skills

As shown in Table 3, both the textbook and its achievement tests overwhelmingly prioritize lower-order cognitive skills, a pattern consistent with previous studies (e.g., Mizbani & Chalack, 2017; Assaly & Smadi, 2015; Lee, 2010). While this emphasis may hinder the development of critical thinking (Erdiana & Panjaitan, 2023; Nair et al., 2019), it is important to consider contextual factors that may justify such design choices. For instance, *Select Readings* targets pre-intermediate learners, whose linguistic proficiency may not yet support sustained engagement with higher-order tasks. In such cases, lower-

[DOR: 20.1001.1.17351634.2024.27.2.2.6]

order activities serve as cognitive scaffolds, enabling learners to build foundational knowledge before progressing to more abstract reasoning (Baghaei et al., 2020).

Additionally, institutional expectations and instructors' pedagogical goals may influence the cognitive orientation of assessments. In contexts where standardized testing is prevalent, educators may prioritize test-aligned instruction, inadvertently reinforcing lower-order skills. Similarly, instructors may opt for comprehension-based tasks to ensure linguistic accessibility, particularly in multilingual or resource-constrained settings (Chavda et al., 2023).

5.5. Differences between the textbook and tests in higher-order thinking

Despite the shared emphasis on lower-order skills, the textbook incorporates nearly twice the proportion of higher-order objectives (17.32%) compared to the achievement tests (8%). This discrepancy likely stems from item format differences, as open-ended questions are more conducive to eliciting analysis, evaluation, and creation (Fajari, 2021). The textbook's inclusion of such tasks suggests an instructional intent to foster deeper engagement, even if this intent is not mirrored in the summative assessments.

6. Conclusion and implications

This study examined the cognitive distribution of learning objectives embedded in *Select Readings* and its associated university achievement tests using Bloom's Revised Taxonomy (2001). The findings revealed a marked misalignment between instructional materials and assessments, a tendency toward lower-order cognitive processing, and a notable neglect of procedural and metacognitive domains.

6.1 .Cognitive misalignment between the textbook and tests

While *Select Readings* encompassed all six cognitive levels of Bloom's taxonomy - Remembering, Understanding, Applying, Analyzing, Evaluating, and Creating - the achievement tests only represented four, omitting Applying and Creating entirely. This omission undermines principles of balanced assessment design (Koksal & Ulum, 2018; Kosasih et al., 2022), potentially restricting learners' opportunity to demonstrate critical reasoning, idea synthesis, and real-world transfer.

[DOR: 20.1001.1.17351634.2024.27.2.2.6]

6.2 .Disproportionate focus on lower-order thinking

Quantitative findings confirmed that both the textbook and tests predominantly targeted lower-order thinking skills. In particular, Understanding factual knowledge dominated the achievement tests (77%), while *Select Readings* maintained a more even split between Remembering and Understanding. Despite this slight improvement in instructional content, higher-order objectives (Analyzing, Evaluating, Creating) accounted for only 17.32% of textbook tasks and a mere 8% of test items. This cognitive skew aligns with prior studies (Freahat & Smadi, 2014; Mizbani & Chalack, 2017), which caution that overemphasis on recall and comprehension impedes deeper learning.

6.3. Neglect of procedural and metacognitive knowledge

Both the textbook and tests showed heavy reliance on factual and conceptual domains, with procedural and metacognitive knowledge entirely absent. This gap is particularly concerning in EFL education, where strategic learning and self-regulation are critical for long-term proficiency (Krathwohl, 2002; Sahragard & Alavi, 2016; Tasnimi, 2020). Without exposure to problem-solving or reflective tasks, students may lack the tools needed to monitor their performance, adapt their strategies, or navigate authentic language challenges.

6.4. Pedagogical recommendations

To ensure that learning experiences promote holistic cognitive development, the following recommendations are proposed:

a. Enhancing Higher-Order Thinking in Textbook Design

Material developers should deliberately integrate more tasks that engage Analyzing, Evaluating, and Creating, particularly within conceptual and procedural knowledge domains.

Evidence from this study suggests that open-ended activities and collaborative formats (e.g., group discussions, critical debates) can enrich higher-order engagement within instructional texts.

b. Improving Cognitive Depth in Summative Assessments

[DOR: 20.1001.1.17351634.2024.27.2.2.6]

Achievement tests must diversify their formats beyond multiple-choice questions to include constructed-response and scenario-based items. These allow students to demonstrate reasoning, argumentation, and synthesis - skills absent from purely recognition-based formats.

Aligning test content with textbook objectives will foster constructive coherence and reinforce students' capacity to transfer knowledge across contexts.

c. Integrating Procedural and Metacognitive Knowledge

Instructional materials and assessments should embed tasks that cultivate procedural fluency and self-awareness. Examples include inquiry-based projects, learning strategy logs, and reflection prompts that encourage students to evaluate their learning approaches.

Evaluation frameworks should evolve to consider metacognitive performance indicators, such as goal-setting behavior, monitoring strategies, and adaptability.

6.5. Final considerations

While these implications offer actionable pathways, their implementation must remain sensitive to contextual factors - including learners' proficiency levels, institutional constraints, and assessment cultures. For example, pre-intermediate learners may require scaffolded progression before engaging with complex cognitive tasks. Similarly, resource limitations or curricular mandates may shape what formats are feasible. Therefore, instructional designers and educators are encouraged to adopt a balanced approach, gradually expanding cognitive complexity while supporting accessibility, equity, and learner agency.

[DOR: 20.1001.1.17351634.2024.27.2.2.6]

Declaration of Conflicting Interests

The authors declare no conflict of interest with respect to the research, authorship, or publication of this study.

References

Anderson, D., & Krathwohl, R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.

Ary, D., Jacobs, L. C., Sorensen, C., & Walker, D. (2019). *Introduction to research in education* (10th ed.). Wadsworth, Cengage Learning.

Assaly, I. R., & Smadi, O. M. (2015). Using Bloom's taxonomy to evaluate the cognitive levels of master class textbook's questions. *English Language Teaching, 8*(5), 100–111. https://doi.org/10.5539/elt.v8n5p100

Astrid, A., Hasanah, A. & Syafryadin, S. (2022). Integrating higher order thinking skills into English language teaching for elementary school students: Teachers' perspectives and challenges. *3L: The Southeast Asian Journal of English Language Studies, 28*(3), 217-230. https://doi.org/10.17576/3l-2022-2803-14

Bachman, L. F. (2013). Ongoing challenges in language assessment. In A. J. Kunnan (Ed.), *The companion to language assessment*. Wiley.

Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice*. Oxford University Press.

Baghaei, S., Bagheri, M. S., & Yamini, M. (2020). Analysis of IELTS and TOEFL reading and listening tests in terms of Revised Bloom's Taxonomy. *Cogent Education, 7*(1), 1–23. https://doi.org/10.1080/2331186X.2020.1770925

Chavda, M., Patel, H. & Bhat, H. (2023). Quality education through writing: Aligning learning objectives in learning materials and question papers using Bloom's taxonomy. *Quality Assurance in Education, 32*(1), 96-110. https://doi.org/10.1108/QAE-03-2023-0045

Divsar, H. & Jafarigohar, M. (2020). Do official curricula reflect critical thinking? The educational objectives of the English literature curricula in Iran in light of Bloom's revised taxonomy. *Iranian Journal of Applied Linguistics, 23*(1), 1-27.

Erdiana, N. and Panjaitan, S. (2023). How is HOTS integrated into the Indonesian high school English textbook? *Studies in English Language and Education, 10*(1), 60-77. https://doi.org/10.24815/siele.v10i1.26052

Fajari, L. E. W. (2021). Open-ended questions to assess critical thinking in Indonesian Elementary School. *International Journal of Instruction, 14*(1), 615-630. https://doi.org/10.29333/iji.2021.14137a

Freahat, N. M., & Smadi, O. M. (2014). Lower-order and higher-order reading questions in secondary and university level EFL textbooks in Jordan. *Theory and Practice in Language Studies, 4*(9), 1804–1813. https://doi.org/10.4304/tpls.4.9.1804-1813

Guo, N., & Yao, Y. (2021). Language demands of textbooks for learning English in Hong Kong: A multi-stratal analysis. *Linguistics and Education, 63*, 100934. https://doi.org/10.1016/j.linged.2021.100934

Hutchinson, T., & Torres, E. (1994). The textbook as an agent of change. *ELT Journal, 48*(4), 315–328. https://doi.org/10.1093/elt/48.4.315

Jovanović, J., Zamecnik, A., Barthakur, A., et al. (2025). Curriculum analytics: Exploring assessment objectives, types, and grades in a study program. *Education and Information Technologies, 30*(1), 4843–4866. https://doi.org/10.1007/s10639-024-12345-6

Kamlasi, I., & Sahan, A. (2018). Descriptive analyses on English test items based on the application of revised Bloom's taxonomy. *Journal of English Language Literature and Teaching, 2*(2), 203–210.

Koksal, D., & Ulum, Ö. G. (2018). Language assessment through Bloom's taxonomy. *Journal of Language and Linguistic Studies, 14*(2), 76–88.

Kosasih, A., Supriyadi, T., Firmansyah, M., & Rahminawati, N. (2022). Cognitive levels in Indonesian EFL textbooks: A Bloom's taxonomy analysis. *Journal of Ethnic and Cultural Studies, 9*(1), 56–76. https://doi.org/10.29333/ejecs/1009

Kowalski, E., M., Koleci, C. & McDonald, K. J.(2024). Changing levels of Bloom's taxonomy in learning objectives and exam questions in first-semester introductory chemistry before and during the adoption of guided inquiry. *Education Sciences, 14*, 943. https://doi.org/10.3390/educsci14090943

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice, 41*(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2

Kuckartz, U. (2025). *Qualitative content analysis: Methods, practice and software.* Stefan Rädiker Publication. https://doi.org/10.4135/9781036212940

Le, H., & Chong, S. L. (2024). The dynamics of critical thinking skills: An exploration from Malaysian and Vietnamese undergraduates. *Thinking Skills and Creativity, 51*, 101445. https://doi.org/10.1016/j.tsc.2023.101445

Lee, L., & Gunersen, E. (2011). *Select readings*. Oxford University Press.

Li, F., & Wang, L. (2024). A study on textbook use and its effects on students' academic performance. *Disciplinary and Interdisciplinary Science Education Research, 6*(4). https://doi.org/10.1186/s43031-024-00089-3

Littlejohn, A. (2011). The analysis of language teaching materials: Inside the Trojan Horse. In B. Tomlinson (Ed.), *Materials development in language teaching* (2nd ed., pp. 179–211). Cambridge University Press.

McDonough, J., & Shaw, C. (2003). *Materials and methods in ELT: A teacher's guide*. Blackwell.

McGrath, I. (2002). *Materials evaluation and design for language teaching*. Edinburgh University Press.

Mizbani, M., Salehi, H., & Tabatabaie, O. (2020). Content evaluation of Iranian EFL textbook Vision 1 based on Bloom's Revised Taxonomy. *International Journal of Foreign Language Teaching and Research, 8*(29), 11–24.

Mizbani, M., & Chalack, A. (2017). Analyzing reading and writing activities of Iranian EFL textbook Prospect 3 through Bloom's Revised Taxonomy. *Journal of Applied Linguistics and Language Research, 4*(2), 13–27.

[DOR: 20.1001.1.17351634.2024.27.2.2.6]

Mukundan, J. (2007). Evaluation of English language textbooks: Some important issues for consideration. *Journal of NELTA, 12*(2), 80–94.

Nair, S. M., Yan, Z. & Wider, W. (2019). Analysis of ESL students' comprehension of LOTS and HOTS questions according to gender. *Humanities and Social Sciences Letters, 7*(2), 74-89. https://doi.org/10.18488/journal.73.2019.72.74.89

Omar, N., Haris, S. S., Hassan, R., Arshad, H., Rahmat, M., Zainal, N. F. A., & Zulkifli, R. (2012). Automated analysis of exam questions according to Bloom's taxonomy. *Procedia - Social and Behavioral Sciences, 59*, 297–303. https://doi.org/10.1016/j.sbspro.2012.09.278

Palmer, E., & Devitt, P. (2006). Constructing multiple-choice questions as a method for learning. *Annals of the Academy of Medicine, 35*(9), 604–608.

Para, S. (2024). A qualitative content analysis exploring gender representation in language textbooks published by Jammu and Kashmir Board of School Education. *Social Sciences and Humanities Open, 10*, 100123. https://doi.org/10.1016/j.ssaho.2024.100123

Quana, P., Shi, Y., Niu, L., & Liu, Y. (2018). Automated estimation of item difficulty for multiple-choice tests: An application of word embedding techniques. *Information Processing & Management, 54*(6), 969–984. https://doi.org/10.1016/j.ipm.2018.07.003

Razmjoo, S. A. (2010). Developing a textbook evaluation scheme for the expanding circle. *Iranian Journal of Applied Language Studies, 2*(1), 121–136.

Razmjoo, A., & Kazempourfard, E. (2012). On the representation of Bloom's Revised Taxonomy in Interchange course books. *The Journal of Teaching Language Skills, 4*(1), 171–204.

Richards, J. C., & Renandya, W. A. (2002). *Methodology in language teaching: An anthology of current practice*. Cambridge University Press.

Roohani, A., Taheri, F., & Poorzangeneh, M. (2013). Evaluating Four Corners textbooks in terms of cognitive processes using Bloom's Revised Taxonomy. *Journal of Research in Applied Linguistics, 4*(2), 51–67.

Salas, A., Tan, M., Andrienko, S., Cengiz, K. & Wisco, J. J. (2024) Reaching the top of Bloom's Taxonomy: An innovative pilot program for preclinical undergraduate and medical school students to create curricula for STEMM outreach/service-learning programs. *Frontiers in Education, 9*, 1446513. https://doi.org/10.3389/feduc.2024.1446513

Sahragard, R., & Alavi, S. Z. (2016). Investigating the predominant levels of learning objectives in general English books. *Journal of English Language Teaching and Learning, 8*(17), 93–114.

Shahmohammadi, S. (2018). Textbook evaluation: Looking at Prospect series through teachers' perspectives. *Research in English Language Pedagogy, 6*(2), 182–204.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4–14. https://doi.org/10.3102/0013189x029007004

Tasnimi, M. (2020). The predictive power of syntactic knowledge, vocabulary breadth and metacognitive strategies for L2 reading fluency. *Iranian Journal of Applied Linguistics, 23*(2), 146-171.

Tomlinson, C. A. (2011). *The differentiated classroom: Responding to the needs of all learners* (2nd ed.). ASCD.

Uskokovic, B., & Schirm, S. (2024). Assessment of interactional competence in L2 German: Integrating an innovative rubric to help language teachers effectively evaluate pragmatics. *System, 125*, 103400. https://doi.org/10.1016/j.system.2024.103400

Van Loi, N. & Thanh, K. (2022). Engaging EFL learners in reading: A text-driven approach to improve reading performance. TESL-EJ, 26(2), 1-18. https://doi.org/10.55593/ej.26102a5

Veeravagu, J., Muthusamy, C., Marimuthu, R., & Subrayan, A. (2010). Using Bloom's Taxonomy to gauge students' reading comprehension performance. *Canadian Social Science, 6*(3), 205–212.

Appendix

Coding Scheme Based on Bloom's Revised Taxonomy

This study employed a coding framework adapted from Bloom's Revised Taxonomy (Anderson & Krathwohl, 2001) to classify learning objectives across two dimensions: the **Knowledge Domain** and the **Cognitive Process Dimension**. The taxonomy provides a systematic structure for evaluating instructional materials and assessments by identifying the type of knowledge targeted and the cognitive operations required. The **Knowledge Domain** includes four categories:

- **Factual Knowledge**: Basic elements such as terminology, specific details, and facts.
- **Conceptual Knowledge**: Interrelationships among concepts, classifications, principles, and theories.
- **Procedural Knowledge**: Methods of inquiry, algorithms, techniques, and criteria for using skills.
- **Metacognitive Knowledge**: Awareness and regulation of one's own cognition, including strategic and self-reflective knowledge.

The **Cognitive Process Dimension** comprises six hierarchical levels:

- **Remembering**: Retrieving relevant knowledge from long-term memory.
- **Understanding**: Constructing meaning from instructional messages.
- **Applying**: Using procedures in new situations.
- **Analyzing**: Breaking material into parts and determining relationships.
- **Evaluating**: Making judgments based on criteria and standards.
- **Creating**: Putting elements together to form a novel, coherent whole.

Each cell in the table below represents a unique combination of knowledge type and cognitive process. Codes such as **A1**, **B1**, and **C1** denote specific intersections (e.g., A1 = Remembering Factual Knowledge).

[DOR: 20.1001.1.17351634.2024.27.2.2.6]

Table A1 *Coding Scheme Based on Bloom's Revised Taxonomy*

Knowledge Domain	Remembering	Understanding	Applying	Analyzing	Evaluating	Creating
Factual Knowledge	A1	B1	C1	D0	E0	F0
Conceptual Knowledge	A2	B2	C2			
Procedural Knowledge	A3	B3	C3			
Metacognitive Knowledge	A4	B4	C4	D4	E4	F4

Note. Codes represent the intersection of cognitive processes and knowledge domains. For example, *C2* refers to **Applying Conceptual Knowledge**, while *E4* refers to **Evaluating Metacognitive Knowledge**. Empty cells indicate combinations not observed or not applicable in the analyzed materials.

[DOR: 20.1001.1.17351634.2024.27.2.2.6]