
| RESEARCH ARTICLE

Cognitive Demand in Higher Education Summative Assessment: A Two-Dimensional Analysis Using Bloom's Revised Taxonomy

Soumia BELKAHLA¹ and Noureddine AZMI²

^{1,2} *Research Lab: Language, Culture & Text: Multidisciplinary Approaches, Faculty of Letters and Human Sciences, Cadi Ayyad University, Marrakech, Morocco*

Corresponding Author: Soumia BELKAHLA, **E-mail,** s.belkahla.ced@uca.ac.ma

| ABSTRACT

This study examines the cognitive and knowledge demands embedded in summative assessment questions within an English Department at a Moroccan public university using Bloom's Revised Taxonomy two-dimensional framework. A corpus of 162 examination questions, drawn from 82 examination papers across multiple assessment genres, was analyzed using deductive qualitative content analysis. Questions were classified according to the interaction between cognitive processes (Remember, Understand, Apply, Analyze, Evaluate, Create) and knowledge types (Factual, Conceptual, Procedural, Metacognitive). Findings indicate that assessment demand was concentrated primarily in Conceptual Knowledge (50.0%), with Understand–Conceptual being the most frequent taxonomy cell (21.0%). Procedural application was also strongly represented, particularly through Apply–Procedural tasks (13.0%). Although higher-order cognitive processes such as Analyze, Evaluate, and Create were present, they occurred less frequently than comprehension-oriented tasks. Metacognitive knowledge was minimally represented (1.2%), suggesting limited explicit attention to reflective or self-regulatory thinking. The study argues that Bloom's Revised Taxonomy provides not only a useful analytical framework for examining cognitive demand but also a practical tool for assessment redesign. Drawing on contemporary discussions surrounding generative AI, equity, and epistemic agency, the article proposes scaffolded assessment structures as a strategy for distributing cognitive demand more deliberately while supporting transparent, equitable, and higher-order forms of disciplinary reasoning in higher-education assessment.

| KEYWORDS

Bloom's Revised Taxonomy, cognitive demand, summative assessment, higher education assessment, assessment design, generative AI, scaffolded assessment, epistemic agency

| ARTICLE INFORMATION

ACCEPTED: 10 May 2026

PUBLISHED: 10 June 2026

DOI: 10.32996/bjtep.2026.5.6.3

1. INTRODUCTION

Assessment is widely recognized as one of the most influential dimensions of teaching and learning in higher education because it shapes what students perceive as valued knowledge and competent performance within a discipline. Research consistently demonstrates that assessment strongly influences students' approaches to learning: tasks emphasizing recall and reproduction tend to encourage surface learning strategies, whereas tasks requiring interpretation, application, evaluation, and synthesis are more likely to promote deeper cognitive engagement and higher-order thinking (Biggs, 1996; Gibbs & Simpson, 2005). Consequently, analyzing assessment tasks themselves provides important insight into the cognitive expectations embedded within higher education practices.

In the Moroccan higher education context, assessment combines continuous assessment with summative examinations, including mid-term and terminal assessments that play a central role in certification and progression decisions (National Charter of Education

and Training [NCET], 2013). Because such examinations are high-stakes and institutionally visible, they provide a particularly revealing site for investigating the kinds of thinking and knowledge that university assessment practices prioritize. This issue is especially significant in English departments, where assessment spans multiple subfields and genres, including literature, linguistics, communication, translation, reading comprehension, and academic writing, each of which may place distinct cognitive demands on students.

Previous research examining assessment through Bloom's Taxonomy has frequently reported a predominance of lower-order cognitive processes, particularly remembering and understanding, alongside comparatively limited engagement with analysis, evaluation, and creation (Crowe et al., 2008; Virranmäki et al., 2020). However, many studies rely primarily on verb-based classification or focus only on the cognitive process dimension, providing limited insight into how cognitive processes interact with different forms of knowledge. Moreover, relatively little research has applied Bloom's Revised Taxonomy in its full two-dimensional form to examine summative assessment practices across multiple assessment genres within Moroccan higher education contexts.

These concerns have become increasingly important in the context of digital transformation and generative artificial intelligence (AI). As AI tools become capable of producing fluent explanations, summaries, and interpretations, assessment tasks that rely primarily on surface-level discussion may become less effective indicators of students' independent reasoning. Recent research therefore emphasizes the need for assessment designs that foreground evidence-based reasoning, evaluative judgment, and transparent justification processes (QAA, 2023; Xia et al., 2024). Although the present study does not investigate AI use directly, this context reinforces the importance of examining the cognitive and knowledge demands embedded in current summative assessment practices.

To address these issues, the present study applies Bloom's Revised Taxonomy (Anderson & Krathwohl, 2001) in its two-dimensional form to analyze the cognitive demand of summative examination questions in a Moroccan university English department. The dataset comprises 162 non-repeated undergraduate examination questions extracted from 86 examination papers across multiple courses and assessment genres. Each question (or sub-question where applicable) was coded according to its dominant cognitive process and knowledge type, enabling a detailed analysis of how cognitive demand is distributed across the taxonomy matrix.

Accordingly, this study addresses the following research questions:

1. Which cognitive processes are most frequently elicited by summative examination questions in the corpus?
2. Which knowledge types are most frequently required?
3. How are cognitive processes and knowledge demands distributed across the 24 cells of Bloom's Revised Taxonomy matrix, and what gaps or imbalances become visible?

2. LITERATURE REVIEW

2.1 Assessment in Higher Education

Research on higher education assessment increasingly emphasizes that assessment practices do more than measure learning outcomes; they shape the forms of knowledge, reasoning, and participation that become institutionally valued (Boud & Falchikov, 2007; Carless, 2015). From the perspective of constructive alignment, assessment tasks should reflect intended learning outcomes and encourage the forms of cognitive engagement that programs aim to develop (Biggs, 1996). Examination questions can therefore be understood as pedagogical and institutional artefacts that reveal the cognitive expectations embedded within assessment systems.

In higher education, assessment is commonly differentiated according to purpose, particularly between formative assessment and summative assessment. Formative assessment aims to support learning during instruction through feedback, guidance, and opportunities for improvement, whereas summative assessment evaluates learning outcomes at designated points for purposes of grading, certification, and progression (Black & William, 1998; Boud & Falchikov, 2007). Relatedly, scholars distinguish between assessment *for* learning, which emphasizes developmental feedback and learner improvement, and assessment *of* learning, which focuses on measuring and certifying achievement (Assessment Reform Group, 2002). Although formative approaches are widely advocated for supporting self-regulated learning, summative examinations remain dominant in many higher education contexts because of their institutional role in accountability and decision-making (Nicol & Macfarlane-Dick, 2006).

In the Moroccan higher education context, assessment combines continuous assessment with summative examinations, including mid-term and terminal assessments, which constitute central components of module evaluation according to the National Charter of Education and Training (NCET, 2013). Because such examinations are high-stakes and institutionally visible, they provide a

particularly important site for investigating the kinds of knowledge and reasoning that higher education systems routinely prioritize and reward.

The educational effects of assessment depend largely on the alignment between intended learning outcomes, teaching activities, and assessment tasks (Biggs, 1996). When programmes aim to develop critical engagement and disciplinary reasoning, but examinations predominantly require recall or low-level comprehension, tensions may emerge between curricular intentions and assessment practices. Analyzing examination questions as curricular artefacts therefore provides insight into whether assessment design is likely to support deeper learning and cognitively demanding forms of engagement (Gibbs & Simpson, 2005).

2.2 Cognitive Demand in Assessment

Research on assessment increasingly focuses on the cognitive demand embedded in tasks, that is, the kinds of thinking students must perform to respond successfully. Cognitive demand concerns whether assessment tasks require students to recall information, interpret meaning, apply procedures, analyze relationships, evaluate alternatives, or generate original responses (Brookhart, 2010). Advanced cognitive performance Higher-order thinking is generally associated with analysis, evaluation, and creation, whereas lower-order thinking is more commonly associated with recall and basic understanding (Resnick, 1987). Tasks requiring more complex cognitive engagement are often considered more likely to support transfer, critical reasoning, and long-term learning goals in higher education (Bransford et al., 2000; Boud & Falchikov, 2006). Recent assessment research further emphasizes that cognitively demanding tasks should require students to justify claims, evaluate evidence, and make disciplinary judgments rather than merely reproduce information (Tai et al., 2018; Xia et al., 2024).

A key methodological issue in the cognitive demand literature is that task verbs alone are insufficient indicators of cognitive complexity. Terms such as “discuss” or “analyze” may elicit very different levels of reasoning depending on whether tasks require evidence, criteria-based judgment, comparison, or justification. Frameworks such as Depth of Knowledge (DOK) therefore emphasize that cognitive demand emerges from the interaction between task requirements and the knowledge students must mobilize, rather than from surface linguistic features alone (Webb, 2002). This insight has motivated the development of more systematic analytical approaches that examine both the cognitive processes required by tasks and the forms of knowledge involved in task performance.

Cognitive demand is also closely related to the type of knowledge assessed. Tasks requiring factual recall may differ substantially from those requiring conceptual understanding, procedural performance, or reflective awareness of one’s own thinking. However, many studies examine cognitive demand without explicitly integrating knowledge types, resulting in potentially oversimplified accounts of what assessment tasks require. Analytical models capable of representing both cognitive processes and knowledge types are therefore particularly valuable for examining examination questions in higher education contexts.

2.3 Bloom’s Revised Taxonomy as an Analytical Framework

Bloom’s Revised Taxonomy conceptualizes learning as the interaction between cognitive processes and knowledge types (Anderson & Krathwohl, 2001). The revised framework reframes cognition as action-oriented processes—Remember, Understand, Apply, Analyze, Evaluate, and Create—thereby emphasizing learning as active cognitive engagement rather than passive acquisition (Figure 1). Unlike the original taxonomy, the revised version also positions *Create* at the highest level of the hierarchy to reflect the complexity involved in generating original products, interpretations, or solutions (Krathwohl, 2002). For more details, please see Figure 1.

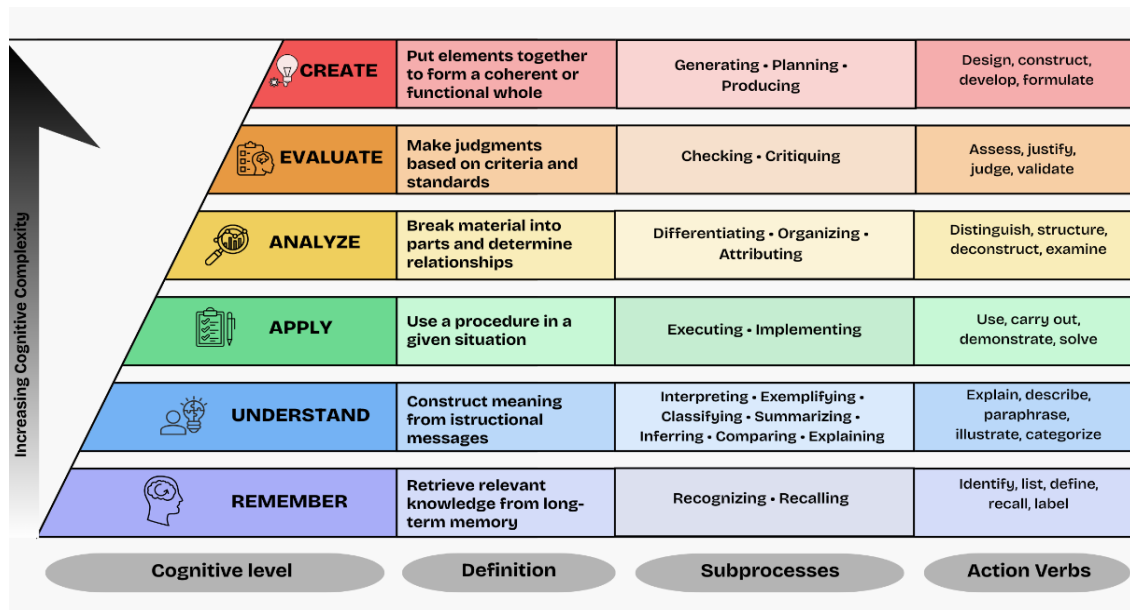


Figure 1: Cognitive Process Taxonomy. Hierarchical and functional classification of cognitive processes adapted from Anderson, L. W., & Krathwohl, D. R. (2001).

A major contribution of the revised taxonomy is the addition of the Knowledge Dimension, which distinguishes among Factual, Conceptual, Procedural, and Metacognitive knowledge (Figure 2). This distinction enables more precise classification of assessment tasks by accounting not only for the type of thinking required, but also for the type of knowledge students must mobilize to respond successfully. For more details, please see Figure 2.

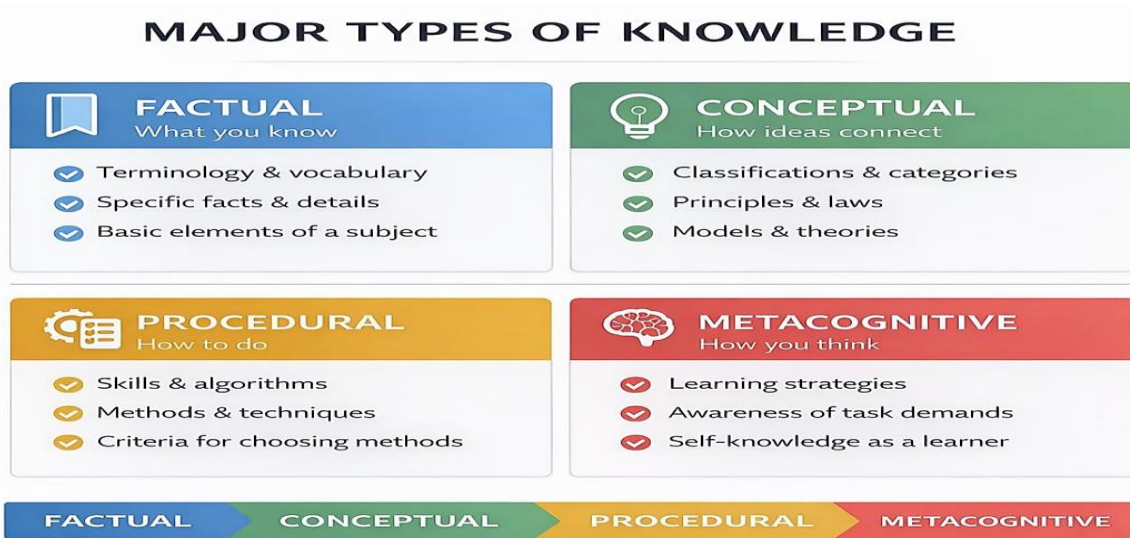


Figure 2: Knowledge dimensions adapted from Anderson, L. W., & Krathwohl, D. R. (2001).

Together, the Cognitive Process and Knowledge dimensions form a two-dimensional taxonomy matrix (Table 1). This structure enables assessment tasks to be classified according to the interaction between cognitive processes (“verb” component) and knowledge types (“noun” component) rather than through directive verbs alone (Krathwohl, 2002). In assessment analysis, the framework is particularly useful for identifying concentrations and absences within assessment systems, such as overrepresentation of understanding-oriented conceptual tasks or limited engagement with metacognitive knowledge. For examination analysis, Bloom’s Revised Taxonomy offers two important advantages: it supports more systematic classification of assessment tasks beyond intuitive or verb-based judgments, and it provides an interpretable structure for examining how cognitive and knowledge demands are distributed across modules, assessment genres, and disciplinary subfields. In the present study, the revised taxonomy serves as

the principal analytical framework for classifying summative examination questions according to their dominant cognitive and knowledge demands. For more details, please see Table 1.

The Knowledge Dimension	The Cognitive Process Dimension					
	1. Remember	2. Understand	3. Apply	4. Analyze	5. Evaluate	6. Create
A. Factual Knowledge						
B. Conceptual Knowledge						
C. Procedural Knowledge						
D. Meta-Cognitive Knowledge						

Table 1: Two-Dimensional Taxonomy Matrix Used for Coding Examination Questions (Adapted from Anderson & Krathwohl, 2001)

Note. The matrix represents the interaction between the Cognitive Process Dimension and the Knowledge Dimension in Bloom's Revised Taxonomy (Anderson & Krathwohl, 2001). Examination questions were coded according to the dominant cognitive process and knowledge type required for an adequate response.

2.4 Empirical Studies Using Bloom's Taxonomy to Analyze Assessments

A substantial body of research has used Bloom's Taxonomy to examine the cognitive demands embedded in assessment tasks across disciplines and educational contexts. A consistent finding across empirical studies is that assessment tasks often cluster within a relatively narrow range of cognitive demands despite curricular emphasis on critical and analytical engagement (Crowe et al., 2008; Muhayimana et al., 2022; Virranmäki et al., 2020). Such patterns raise concerns about whether assessment practices adequately support complex reasoning, interpretive judgment, and transferable learning.

Empirical research also highlights methodological limitations in verb-based approaches to cognitive classification, which may oversimplify assessment demand and obscure the interaction between cognitive processes and knowledge types (Krathwohl, 2002; Webb, 2002). Studies applying Bloom's Revised Taxonomy in its two-dimensional form demonstrate that assessment systems frequently privilege understanding-oriented conceptual tasks while giving comparatively limited attention to evaluative, creative, or metacognitive engagement (Virranmäki et al., 2020). Recent papers further argue that cognitively demanding assessments should require students to justify claims, evaluate evidence, and make criteria-based disciplinary judgments rather than merely reproduce information (Tai et al., 2018; Xia et al., 2024).

Despite growing use of Bloom's Revised Taxonomy, relatively few studies apply the full two-dimensional framework across multiple assessment genres within the same institutional context over extended periods of time. Research examining cognitive demand in Moroccan higher education—and particularly within English Departments—also remains limited. The present study addresses these gaps by applying Bloom's Revised Taxonomy to a longitudinal corpus of summative examination questions across multiple English Department assessment genres.

2.5 Assessment in the Era of Generative AI, Equity, and Epistemic Agency

Recent developments in generative artificial intelligence (AI) have intensified debates concerning what assessment should measure and how assessment tasks can validly demonstrate student learning. As generative AI tools become increasingly capable of producing fluent summaries, explanations, and argumentative responses, concerns have emerged regarding the extent to which conventional examinations can distinguish between routine textual generation and independent disciplinary reasoning (Xia et al., 2024). Recent assessment research therefore emphasizes the need for tasks that foreground accountable reasoning, evaluative judgment, and transparent decision-making processes rather than routine reproduction of information (Bearman et al., 2024; QAA, 2023). Emerging evidence from higher education contexts further suggests that the widespread availability of generative AI is reshaping institutional approaches to assessment design and validity (Perkins et al., 2025).

Within this context, questions of equity and fairness become central to assessment design. Inclusive assessment research argues that assessment practices may either widen or reduce inequities depending on transparency, accessibility, and the range of legitimate

ways students are permitted to demonstrate learning (Bain, 2023). Assessment tasks relying heavily on implicit expectations or narrowly defined forms of academic discourse may disadvantage some students despite their possession of relevant understanding and disciplinary capabilities. Conversely, assessment designs that clarify evaluative criteria, provide structured entry points, and recognize multiple forms of evidence may better support equitable participation and more inclusive demonstrations of competence. International policy guidance similarly highlights the importance of human-centered approaches to AI integration that address risks related to bias, misinformation, and unequal access in technology-mediated educational environments (UNESCO, 2023).

These concerns are closely connected to the concept of *epistemic agency*, understood as students' capacity to act as knowers who justify claims, evaluate evidence, and take responsibility for disciplinary judgments. Recent research suggests that AI-mediated learning environments may reshape epistemic relations by influencing how knowledge is produced, validated, and trusted within educational settings (Nieminen & Ketonen, 2024). In assessment contexts, tasks requiring evidence selection, criteria-based judgment, and explicit rationale statements can strengthen epistemic agency by making reasoning processes visible and assessable. Relatedly, the concept of *evaluative judgment*—the capability to judge the quality of work—has increasingly been framed as an essential graduate attribute that assessment should actively cultivate rather than merely assume (Boud, 2000; Tai et al., 2018). Recent discussions of AI-responsive assessment similarly emphasize the importance of helping students critically evaluate the credibility, appropriateness, and limitations of AI-generated outputs (Bearman et al., 2024).

One practical implication emerging from these discussions is the growing interest in scaffolded assessment structures, including scaffolded examinations. Scaffolding involves providing structured support during initial stages of learning or task performance and gradually reducing support as learners assume greater independence (Wood et al., 1976). Cognitive apprenticeship extends this principle through modeling, coaching, and guided participation designed to support the development of disciplinary expertise (Collins et al., 1989). Applied to examinations, scaffolded assessment can distribute cognitive demand more deliberately across levels, beginning with accessible tasks that establish shared footing and progress toward independent analysis, evaluation, and creation. Such structures may support fairness by reducing cognitive overload and clarifying expectations while still enabling the assessment of advanced disciplinary reasoning and higher-order cognition. Contemporary discussions of AI-responsive assessment similarly advocate process-oriented designs that make reasoning, justification, and evaluative decision-making more visible within assessment tasks (Perkins et al., 2025).

2.6 Positioning the Present Study

Collectively, the literature reviewed above demonstrates that assessment is not merely a mechanism for measuring achievement, but a central pedagogical practice that shapes the forms of knowledge, reasoning, and participation that become institutionally valued in higher education (Biggs, 1996; Gibbs & Simpson, 2005). Research employing Bloom's Taxonomy has further shown the value of systematically examining assessment tasks to identify patterns in cognitive demand, while simultaneously highlighting methodological limitations associated with verb-based classification approaches that fail to account for the interaction between cognitive processes and knowledge types (Anderson & Krathwohl, 2001; Krathwohl, 2002; Webb, 2002). At the same time, emerging research on generative AI and assessment increasingly argues that assessment validity now depends less on students' ability to produce fluent textual outputs and more on the extent to which assessment tasks elicit accountable reasoning, evaluative judgment, evidence use, and transparent decision-making processes (QAA, 2023; Xia et al., 2024). Parallel developments in inclusive assessment and epistemic agency studies similarly emphasize the importance of assessment designs that support equitable participation and position students as active disciplinary knowers capable of justifying interpretations and critically evaluating evidence (Bain, 2023; Nieminen & Ketonen, 2024).

Despite these developments, relatively limited research has applied Bloom's Revised Taxonomy in its full two-dimensional form to examine summative assessment practices across multiple assessment genres within English Departments, particularly in Moroccan higher education contexts. Existing studies also rarely connect analyses of cognitive demand to broader contemporary concerns surrounding assessment validity, equity, epistemic agency, and AI-mediated educational contexts. Addressing these gaps, the present study applies Bloom's Revised Taxonomy as both an analytical and reform-oriented framework to examine the distribution of cognitive processes and knowledge demands across a longitudinal corpus of English Department summative examination questions. By mapping assessment tasks across the taxonomy matrix, the study identifies patterns of emphasis, imbalance, and omission in current assessment practices while also demonstrating how the two-dimensional framework can inform more intentional and cognitively balanced assessment design. In doing so, the study contributes not only to research on cognitive demand in higher education assessment but also to ongoing discussions concerning equitable, future-oriented, and AI-responsive approaches to assessment reform.

The following section outlines the methodological procedures used to operationalize this analytical framework and systematically classify examination questions across the taxonomy matrix.

3. METHODOLOGY

3.1 Research Design

This study employed **deductive qualitative content analysis** to examine the cognitive and knowledge demands embedded in higher education assessment questions. Bloom's Revised Taxonomy (Anderson & Krathwohl, 2001) was used as an *a priori* analytical framework, enabling systematic classification of questions according to two dimensions: **cognitive process** and **knowledge type**. This design was appropriate because the study aimed to map assessment demand through a theory-driven coding scheme rather than generate categories inductively.

3.2 Dataset and Context

The dataset consisted of summative undergraduate examination questions collected from the English Department of a Moroccan public university. The corpus included examination papers from multiple courses and assessment genres, including linguistics, literature, communication, translation, reading comprehension, and writing/composition. All materials were drawn from the same academic department across several academic years, allowing the study to examine recurring patterns in assessment design within a shared institutional context.

The final analytical corpus comprised **162 non-repeated examination questions** extracted from **86 examination papers**. These questions represented multiple academic levels, semesters, and formats, including essay-based tasks, short-answer questions, text-based interpretation tasks, translation exercises, and performance-oriented prompts. This diversity allowed for analysis across assessment genres while maintaining institutional consistency.

3.3 Unit of Analysis

The unit of analysis was a single assessment question or sub-question. Where an examination item contained multiple parts (e.g., a/b/c), each part was treated as a separate unit when it required a distinct student response or cognitive operation. Where sub-parts were integrated and required a single holistic response, the item was coded according to its dominant cognitive and knowledge demand.

3.4 Analytical Framework

Each unit of analysis was coded using Bloom's Revised Taxonomy two-dimensional framework, comprising:

- **Cognitive Process Dimension:** Remember, Understand, Apply, Analyze, Evaluate, and Create
- **Knowledge Dimension:** Factual, Conceptual, Procedural, and Metacognitive Knowledge

Coding involved locating each unit within the **24-cell taxonomy matrix**. Each question was assigned to one dominant cognitive process code and one dominant knowledge type code. The cognitive process code reflected the main operation required for an adequate response, while the knowledge type code reflected the primary form of knowledge students were expected to mobilize. Classification was based not only on directive verbs but also on the evidentiary expectations of the prompt, such as whether the task required explanation, application of a procedure, comparison, justification, critique, or original production.

3.5 Coding Procedure

Coding proceeded in three stages. First, examination questions were segmented and entered into an analysis table. Second, each unit was assigned two codes: one for the cognitive process dimension and one for the knowledge dimension, using operational definitions aligned with Bloom's Revised Taxonomy. Third, coded items were reviewed iteratively to ensure consistency across the dataset.

Borderline cases, particularly prompts using broad verbs such as *discuss*, *comment on*, or *analyze*, were resolved by considering the level of reasoning required for an adequate answer. For example, prompts requiring justification, criteria-based judgment, comparison, decomposition, or original production were coded at higher cognitive levels than prompts requiring explanation or description only.

3.6 Trustworthiness of Coding

To strengthen analytical trustworthiness, coding decisions were guided by explicit operational definitions and applied consistently across the corpus. Ambiguous cases were revisited during iterative review, and the final coding emphasized the dominant demand required by each question rather than relying solely on surface wording. This procedure helped reduce overdependence on command verbs and supported a more valid interpretation of cognitive demand.

3.7 Data Analysis

After coding, frequencies and percentages were calculated for each cognitive process, each knowledge type, and each cell in the 6 × 4 taxonomy matrix. This allowed the study to identify dominant categories, underrepresented areas, and patterns of concentration across the assessment corpus. The quantitative summaries were then interpreted qualitatively in relation to assessment design, higher-order thinking, and the distribution of cognitive demand across English Department summative assessment practices.

4. RESULTS

A total of 162 assessment questions were coded using Bloom's Revised Taxonomy two-dimensional framework. Across the Knowledge Dimension, **Conceptual Knowledge** was the most frequently represented category (n = 81, 50.0%), followed by **Procedural Knowledge** (n = 47, 29.0%) and **Factual Knowledge** (n = 32, 19.8%). **Metacognitive Knowledge** was minimally represented (n = 2, 1.2%).

Across the Cognitive Process Dimension, **Understand** was the most frequent category (n = 43, 26.5%). **Remember** and **Apply** were equally represented (n = 27, 16.7% each), followed by **Analyze** (n = 23, 14.2%), **Create** (n = 22, 13.6%), and **Evaluate** (n = 20, 12.3%). Overall, the distribution suggests a concentration of assessment demand within comprehension-oriented and procedural tasks.

At the cell level, the most prevalent category was **Understand-Conceptual** (n = 34, 21.0%), indicating that assessment questions frequently required students to explain, interpret, or demonstrate understanding of conceptual relationships and meanings. The second most frequent category was **Apply-Procedural** (n = 21, 13.0%), reflecting substantial emphasis on the implementation of learned procedures and conventions. **Remember-Factual** was also prominent (n = 20, 12.3%), demonstrating continued reliance on recall of factual information and terminology.

Higher-order cognitive demands were present but less prominent overall. **Analyze-Conceptual** accounted for 13 questions (8.0%), while **Evaluate-Conceptual** accounted for 16 questions (9.9%), indicating some attention to analytical reasoning and criteria-based judgment. Creation-oriented tasks were distributed across **Create-Procedural** (n = 11, 6.8%) and **Create-Conceptual** (n = 9, 5.6%), suggesting occasional opportunities for original production and synthesis.

Metacognitive assessment was notably limited. Only two questions were classified within the metacognitive dimension: **Evaluate-Metacognitive** (n = 1, 0.6%) and **Create-Metacognitive** (n = 1, 0.6%). No questions were coded within **Remember-Metacognitive**, **Understand-Metacognitive**, **Apply-Metacognitive**, or **Analyze-Metacognitive** categories. For more details, please see Table 2.

Cognitive Process Knowledge type	Factual	Conceptual	Procedural	Metacognitive	Row total
Remember	20 (12.3%)	5 (3.1%)	2 (1.2%)	0 (0.0%)	27 (16.7%)
Understand	5 (3.1%)	34 (21.0%)	4 (2.5%)	0 (0.0%)	43 (26.5%)
Apply	2 (1.2%)	4 (2.5%)	21 (13.0%)	0 (0.0%)	27 (16.7%)
Analyze	3 (1.9%)	13 (8.0%)	7 (4.3%)	0 (0.0%)	23 (14.2%)
Evaluate	1 (0.6%)	16 (9.9%)	2 (1.2%)	1 (0.6%)	20 (12.3%)
Create	1 (0.6%)	9 (5.6%)	11 (6.8%)	1 (0.6%)	22 (13.6%)
Column total	32 (19.8%)	81 (50.0%)	47 (29.0%)	2 (1.2%)	162 (100.0%)

Table 2: Bloom's Revised Taxonomy Two-Dimensional Frequency Matrix (N = 162)
Note. Cell entries represent frequencies and percentages of the full corpus (N = 162).

Overall, the findings indicate that assessment demand was concentrated primarily within conceptual understanding and procedural application categories, while metacognitive engagement remained minimal across the corpus. Figure 3 visually reinforces this concentration within Understand–Conceptual and Apply–Procedural categories. The following Discussion section examines the implications of this distribution for higher-order cognition, assessment validity, equity, and assessment redesign in contemporary higher education contexts.

5. DISCUSSION

This study used Bloom’s Revised Taxonomy two-dimensional framework to audit the cognitive and knowledge demands of higher education assessment questions in an English Department corpus (N = 162). The distribution in Table 1 indicates that assessment emphasized Conceptual knowledge (50.0%), with the single most frequent cell being Understand–Conceptual (21.0%), alongside substantial attention to procedural performance (e.g., Apply–Procedural, 13.0%). Metacognitive demands were rare (1.2% overall). These patterns offer more than a descriptive profile; they identify where assessment currently concentrates academic value and, therefore, where reform is most needed if departments aim to cultivate higher-order thinking, equitable participation, and student epistemic agency in the context of digital transformation (Anderson & Krathwohl, 2001; Krathwohl, 2002; Biggs, 1996).

Assessment reform in the AI and digital transformation moment

The predominance of understanding-oriented conceptual tasks suggests that many assessments rely on demonstrations of comprehension through explanation and interpretation. In an AI-rich environment, where generative tools can readily produce fluent summaries and plausible interpretations, assessments dominated by “explain/discuss” prompts risk becoming less diagnostic of students’ independent reasoning unless they are redesigned to require stronger forms of evidence (Xia et al., 2024).

This does not diminish the importance of conceptual understanding in English studies; rather, it underscores the need to re-specify what counts as evidence of understanding. For example, interpretation prompts can be strengthened by requiring students to justify claims against explicit criteria, evaluate alternative readings, or tie interpretations to specific textual, linguistic, or rhetorical evidence. Such design shifts foreground reasoning and warranting, increasing validity when surface fluency is no longer a reliable indicator of learning (Kane, 2013; Sadler, 1989).

Equity, fairness, and epistemic agency as assessment outcomes

The corpus’s concentration on conventional explanatory prose—particularly Understand–Conceptual tasks—has direct equity implications. When the dominant evidentiary form is extended written explanation, students less familiar with dominant academic language practices may be disadvantaged despite possessing genuine disciplinary understanding. Rebalancing the taxonomy profile to include structured analytic moves, criterion-based judgments, and constrained productions can widen the range of legitimate evidence for learning while making expectations more transparent (Bain, 2023).

The near absence of metacognitive demand is particularly relevant here: when assessments rarely ask students to articulate strategy, rationale, or limitations, they also limit students’ epistemic agency—their capacity to act as knowers who can justify decisions, monitor reasoning, and take responsibility for claims (Nieminen & Ketonen, 2024). In digitally transformed contexts, epistemic agency includes the ability to evaluate information quality, verify claims, and transparently justify choices; making these practices assessable can support both equity (through explicit criteria) and integrity (through process visibility) (UNESCO, 2023).

Higher-order thinking: present but not systematically rewarded

Although higher-order processes (Analyze, Evaluate, Create) were present in the corpus, they occurred less frequently than Understand and Apply. This suggests that students may have intermittent opportunities to demonstrate advanced reasoning, but that such performance may not be consistently elicited or rewarded across assessments. Because assessment strongly shapes student learning behavior, limited and uneven inclusion of higher-order tasks can signal that comprehension and routine performance are sufficient for success. A reform implication is to move toward deliberate assessment blueprinting at module and program levels, ensuring repeated and developmental engagement with analysis, evaluation, and synthesis—particularly in advanced years and capstone modules (Biggs, 1996; Boud & Falchikov, 2006).

Using Bloom’s 2D framework as a tool for reimagining assessment

A key contribution of this study is to position the two-dimensional taxonomy not only as a classificatory device but as a practical design tool for assessment reform. The matrix makes visible patterns that are otherwise difficult to detect (i.e., such as concentration in Understand–Conceptual and minimal metacognitive assessment) and supports intentional rebalancing aligned with stated educational goals (Anderson & Krathwohl, 2001; Krathwohl, 2002).

Used prospectively, the framework can guide (a) the redesign of prompts to elicit higher-order reasoning, (b) the diversification of evidentiary forms students can use to demonstrate learning, and (c) the incorporation of explicit criteria that enhance transparency and fairness. In this sense, the method supports reimagining assessment as a structured opportunity for students to enact disciplinary thinking—interpretive, analytic, critical, and creative—rather than primarily reproducing conventional explanations (Sadler, 1989; Nieminen & Ketonen, 2024).

Scaffolded exams as a reform strategy for higher-order cognition and fairness

To operationalize these implications, **scaffolded exam design** offers a practical strategy for distributing cognitive demand more deliberately across Bloom's categories while supporting equitable access to higher-order assessment tasks. Rather than concentrating assessment primarily within comprehension-oriented prompts, scaffolded examinations can sequence tasks progressively from foundational understanding toward more complex forms of disciplinary reasoning. As summarized in Table 3, such designs may begin with **accessible entry tasks** that establish shared conceptual footing (**Remember/Understand**), transition to **guided procedural application (Apply)**, and culminate in **analytical interpretation, criteria-based judgment, and constrained synthesis (Analyze/Evaluate/Create)**. This staged progression makes higher-order expectations more explicit and progressively attainable while reducing abrupt shifts in cognitive demand.

A further advantage of scaffolded assessment is that it can preserve **disciplinary authenticity** while supporting cognitive progression. In English studies, a single shared stimulus—such as a literary extract, transcript, editorial text, advertisement, linguistic dataset, or student writing sample—can anchor multiple stages of questioning. This enables adaptation across literature, linguistics, rhetoric/composition, translation, cultural studies, and teacher-education modules while maintaining coherence across the assessment task. Because authentic disciplinary questions often activate multiple taxonomy categories simultaneously, Table 3 identifies the dominant cognitive focus associated with each stage rather than implying rigid separation among taxonomy cells.

In **AI-mediated educational contexts**, scaffolded examinations also provide a principled way to shift assessment emphasis from routine textual generation toward **accountable disciplinary reasoning**. Later-stage prompts that require evidence selection, explicit criteria use, and justified interpretation may better differentiate disciplinary judgment from generic AI-supported output (Xia et al., 2024). Incorporating brief rationale components (i.e., requiring students to explain interpretive choices, justify evidence selection, or acknowledge limitations) can make reasoning processes more visible without substantially increasing marking burden. Such approaches may strengthen assessment validity and fairness while also cultivating evaluative judgment as an essential graduate capability (Boud, 2000; Tai et al., 2018).

A concise illustrative progression is presented in Table 3, while a more detailed example of a scaffolded examination blueprint is provided in Appendix A. For more details, please see Table 3 and Appendix A.

Stage	Dominant Bloom Focus	Assessment Purpose
Entry tasks	Remember / Understand	Establish shared conceptual footing
Guided application	Apply	Operationalize disciplinary procedures
Analytical interpretation	Analyze	Examine patterns, relationships, and evidence
Judgment tasks	Evaluate	Make criteria-based judgments
Constrained synthesis	Create	Produce justified disciplinary responses

Table 3: Illustrative Scaffolded Examination Progression Aligned with Bloom's Revised Taxonomy.

Limitations and future research

This study examined the intended cognitive and knowledge demands embedded in summative examination questions rather than students' actual cognitive performance or the ways assessment criteria were applied in practice. Although Bloom's Revised Taxonomy provided a systematic analytical framework, coding inevitably involved interpretive judgment, particularly in cases where directive verbs (e.g., *discuss*) could imply multiple cognitive demands depending on evidentiary expectations. In addition, the corpus was drawn from a single English Department within one Moroccan public university and therefore should not be interpreted as representative of all higher education contexts. Finally, while the article discusses implications for AI-responsive, equitable, and scaffolded assessment design, these implications remain conceptual rather than empirically tested. Future research could triangulate

taxonomy-based coding with marking rubrics, student scripts, examiner perspectives, and comparative cross-disciplinary analyses to examine how cognitive demand is enacted, assessed, and experienced in practice.

6. CONCLUSION

This study applied Bloom's Revised Taxonomy two-dimensional framework to investigate the cognitive processes and knowledge demands embedded in an English Department assessment corpus (N = 162). The findings show a clear emphasis on **Conceptual knowledge** and **understanding-oriented tasks**, with **Understand-Conceptual** as the most frequent cell, alongside substantial assessment of **procedural performance** (notably **Apply-Procedural**). In contrast, **metacognitive demands** were rarely made visible as assessable outcomes. Taken together, the results suggest that current assessment practices prioritize comprehension and execution, while offering fewer structured opportunities for students to demonstrate sustained higher-order reasoning and explicit self-regulation.

In the context of AI and digital transformation, these patterns have heightened significance. When fluent explanation and routine text production can be readily supported by generative tools, the validity of assessment increasingly depends on tasks that elicit justified judgment, evidence-based reasoning, and transparent decision-making. Framed as a reform agenda, the taxonomy review functions not only as a descriptive account but as a diagnostic tool for identifying where assessment can better support equity, fairness, and epistemic agency, i.e., broadening how students can demonstrate learning and making expectations more explicit.

A key practical implication is the value of scaffolded exam design as a means of rebalancing cognitive demand across Bloom's categories while maintaining inclusive access. By sequencing items from entry-level tasks (Remember/Understand) through guided application (Apply) and toward independent analysis, evaluation, and constrained creation (Analyze/Evaluate/Create), scaffolded exams can both reduce hidden rules and ensure that higher-order thinking is explicitly assessed and rewarded. Future work should triangulate prompt-level analyses with rubrics and student scripts and examine whether taxonomy-informed redesign—including scaffolded structures and AI-aware evidentiary requirements—improves the fairness, integrity, and developmental impact of assessment in English studies.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

REFERENCES

- [1]. Amer, A. (2006). Reflections on Bloom's revised taxonomy. *Electronic Journal of Research in Educational Psychology*, 4(1), 213–230.
- [2]. Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- [3]. Assessment Reform Group. (2002). *Assessment for learning: 10 principles*. University of Cambridge School of Education.
- [4]. Bain, K. (2023). *Inclusive assessment in higher education: Designing for equity and participation*. Routledge.
- [5]. Bearman, M., Tai, J., Dawson, P., Boud, D., & Ajjawi, R. (2024). Developing evaluative judgement for a time of generative artificial intelligence. *Assessment & Evaluation in Higher Education*, 49(6), 893–905. <https://doi.org/10.1080/02602938.2024.2335321>
- [6]. Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education*, 32(3), 347–364. <https://doi.org/10.1007/BF00138871>
- [7]. Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74. <https://doi.org/10.1080/0969595980050102>
- [8]. Boud, D. (2000). Sustainable assessment: Rethinking assessment for the learning society. *Studies in Continuing Education*, 22(2), 151–167. <https://doi.org/10.1080/713695728>
- [9]. Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education*, 31(4), 399–413. <https://doi.org/10.1080/02602930600679050>
- [10]. Boud, D., & Falchikov, N. (2007). *Rethinking assessment in higher education: Learning for the longer term*. Routledge.
- [11]. Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). *How people learn: Brain, mind, experience, and school*. National Academy Press.
- [12]. Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. ASCD.
- [13]. Carless, D. (2015). *Excellence in university assessment: Learning from award-winning practice*. Routledge.
- [14]. Chen, M. (2007). *Education nation: Six leading edges of innovation in our schools*. Jossey-Bass.

- [15]. Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 453–494). Erlbaum.
- [16]. Crowe, A., Dirks, C., & Wenderoth, M. P. (2008). Biology in Bloom: Implementing Bloom's taxonomy to enhance student learning in biology. *CBE—Life Sciences Education*, 7(4), 368–381. <https://doi.org/10.1187/cbe.08-05-0024>
- [17]. Fletcher, R. B., Meyer, L. H., Anderson, H., Johnston, P., & Rees, M. (2012). Faculty and students' conceptions of assessment in higher education. *Higher Education*, 64(1), 119–133. <https://doi.org/10.1007/s10734-011-9484-1>
- [18]. Gibbs, G., & Simpson, C. (2005). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education*, 1, 3–31.
- [19]. Kane, M. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1–73. <https://doi.org/10.1111/jedm.12000>
- [20]. 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
- [21]. Litosseliti, L. (2010). *Research methods in linguistics*. Continuum.
- [22]. Muhayimana, J., Nsanganw'abandi, P., & Uworwabayeho, A. (2022). Analysis of examination questions using Bloom's taxonomy in higher education. *International Journal of Educational Research Review*, 7(2), 145–156.
- [23]. National Charter of Education and Training (NCET). (2013). *National Charter of Education and Training*. Ministry of National Education, Morocco.
- [24]. Nenty, H. J., Adedoyin, O. O., Odili, J. N., & Major, T. E. (2007). Evaluation of educational assessment systems in selected African countries. *Educational Research and Review*, 2(7), 170–177.
- [25]. Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199–218. <https://doi.org/10.1080/03075070600572090>
- [26]. Nieminen, J. H., & Ketonen, L. (2024). Epistemic agency: A link between assessment, knowledge and society. *Higher Education*, 88(2), 777–794. <https://doi.org/10.1007/s10734-023-01142-5>
- [27]. OECD. (2008). *Assessment for learning: Formative assessment*. OECD Publishing.
- [28]. Perkins, G. (2026). Resilient assessment in the age of AI: authentic design and the case for verbal examinations in business education. *Assessment & Evaluation in Higher Education*, 1–16. <https://doi.org/10.1080/02602938.2026.2644516>
- [29]. Quality Assurance Agency for Higher Education. (2023). *Reconsidering assessment for the ChatGPT era*. QAA.
- [30]. Resnick, L. B. (1987). *Education and learning to think*. National Academy Press.
- [31]. Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18(2), 119–144. <https://doi.org/10.1007/BF00117714>
- [32]. Tai, J., Ajjawi, R., Boud, D., Dawson, P., & Panadero, E. (2018). Developing evaluative judgment: Enabling students to make decisions about the quality of work. *Higher Education*, 76(3), 467–481. <https://doi.org/10.1007/s10734-017-0220-3>
- [33]. UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO Publishing.
- [34]. Virranmäki, E., Valta-Hulkkonen, K., & Pellikka, I. (2020). Geography tests in the Finnish matriculation examination in paper and digital forms—An analysis of questions based on revised Bloom's taxonomy. *Studies in Educational Evaluation*, 66, Article 100896. <https://doi.org/10.1016/j.stueduc.2020.100896>
- [35]. Webb, N. L. (2002). *Depth-of-knowledge levels for four content areas*. Wisconsin Center for Education Research.
- [36]. Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, 17(2), 89–100. <https://doi.org/10.1111/j.1469-7610.1976.tb00381.x>
- [37]. Xia, Q., Weng, X., Ouyang, F., Lin, T.-J., & Chiu, T. K. F. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education*, 21(1), 1–34. <https://doi.org/10.1186/s41239-024-00434-1>

Appendix A

Table 3: *Scaffolded exam blueprint for English Department summative exams mapped to Bloom's Revised Taxonomy (2D) Moroccan grading scale (Total = 20)*

SECTION	MARKS (e.g., .../20)	PRIMARY BLOOM 2D CELL(S)	TYPICAL TASK (ADAPTABLE)	SUGGESTED VERBS (DISCIPLINE- NEUTRAL)	EVIDENCE EXPECTED (WHAT MARKERS LOOK FOR)	SCAFFOLDI NG LEVEL
A1. Foundations : key terms & features	2	Remember– Factual	Define core terms or identify/label specified features in the stimulus (e.g., imagery, deixis, cohesion, ethos/pathos/logos).	define, list, label, name, identify, match, select, recall, state, recognize	Accurate definitions/identifications; correct terminology use.	High (clear prompts; bounded answers)
A2. Comprehension: meaning/function	2	Understand– Conceptual	Summarize the main idea/theme/argument and explain purpose or effect (≈150–200 words).	explain, describe, summarize, paraphrase, illustrate, classify, interpret, clarify, exemplify	Coherent explanation; accurate interpretation; relevant reference to stimulus.	High– moderate (structure cues; bounded length)
B1. Apply a provided framework	3	Apply– Procedural	Apply an analytic grid/checklist (e.g., rhetorical moves, discourse features, narrative elements, linguistic categories).	apply, use, implement, complete, demonstrate, employ, carry out, follow, operationalize	Correct application of steps; accurate classification; brief functional comments.	Moderate (grid provided; may include one worked example)
B2. Controlled performance task	2	Apply– Procedural (sometimes Understand– Procedural)	Paraphrase, rewrite for audience/purpose, annotate, transform forms (e.g., register shift; cohesion improvement; translation micro-task).	rewrite, transform, translate, edit, revise, correct, convert, adapt, organize, format	Appropriate execution of conventions; alignment to constraints (audience/purpose); correctness.	Moderate (constraints specified; bounded task)
C1. Pattern-based analysis	3	Analyze– Conceptual	Identify patterns and explain relationships (how features work together to produce meaning/positioning).	analyze, differentiate, distinguish, examine, deconstruct, compare, contrast, map, break down, connect	Analytic claims; evidence selection; explanation of relationships (not just listing).	Low– moderate (students choose evidence)
C2. Lens-based analysis	3	Analyze– Conceptual / Analyze– Procedural	Apply one theoretical/critical lens (e.g., postcolonial, feminist, genre/discourse, sociolinguistic) to the stimulus.	interpret, contextualize, interrogate, unpack, problematize, relate, diagnose, analyze (through...), apply (lens)	Accurate lens use; warranted interpretation; evidence-linked reasoning; coherence.	Low (support fades; student structures response)

D. Criteria-based evaluation	3	Evaluate– Conceptual	Evaluate effectiveness/credibility/ethics or competing interpretations using stated criteria.	evaluate, assess, judge, justify, critique, appraise, defend, recommend, weigh, prioritize	Explicit criteria use; justified judgment; evidence aligned to criteria.	Low (criteria given; reasoning demand high)
E. Synthesis/production (capstone)	2	Create– Conceptual/Procedural	Produce an original artifact (mini-argument, proposal abstract, speech opening, alternative reading plan, short creative rewrite with constraints).	create, design, develop, formulate, propose, compose, produce, generate, devise, construct (open)	Originality + fit to constraints; coherence; optional brief rationale.	Low (open production; minimal supports)