JRIIE Journal of Research Innovation and Implications in Education
Center for Research Implications & Practice

Website: www.jriiejournal.com ISSN 2520-7504 (Online) Vol.10, Iss.3, 2026 (pp. 3111 - 3124)

Evaluating the Capacity of Biology Multiple-Choice Questions to Assess Higher-order Thinking Skills in Secondary Schools: A Case Study of Ruhango District, Rwanda

Dushimimana Emmanuel and Jean de Dieu Habiyaremye
African Centre of Excellence for Innovative Teaching and Learning Mathematics and Science (ACEITLMS), College of Education, University of Rwanda
Email: emmadushi1986@gmail.com

Abstract: *In Rwanda, the Competence-Based Curriculum (CBC) requires classroom assessment to measure Higher-Order Thinking Skills (HOTS), yet doubts persist about the capacity of Multiple-Choice Questions (MCQs) to assess HOTS in Biology. This study evaluated the extent to which Biology MCQs used in secondary schools in Ruhango District, Rwanda, assess HOTS, examined teachers' knowledge, perceptions, and practices regarding HOTS-oriented MCQs, and identified factors influencing their design. A convergent mixed-methods case study design was adopted. Data were drawn from document analysis of 681 Biology MCQs sampled from ten secondary schools, questionnaires administered to 16 Biology teachers, and semi-structured interviews with 16 school leaders (7 Directors of Studies and 9 Headteachers). Quantitative data were analyzed descriptively, while qualitative data were analyzed thematically. Findings showed that 98.7% of MCQs assessed Lower-Order Thinking Skills, while only 1.3% assessed HOTS, and no item assessed creating. Teachers demonstrated adequate theoretical knowledge of HOTS but limited confidence in constructing HOTS-oriented items, constrained by inadequate training, large classes, limited resources, and time pressure. The study concludes that Biology MCQs in Ruhango District do not adequately reflect CBC expectations and recommends targeted teacher training, resource provision, and assessment moderation to strengthen HOTS assessment.*

Keywords*: Multiple Choice Questions, Higher-Order Thinking Skills, Bloom's Revised Taxonomy, Competence-Based Curriculum, Biology Assessment, Ruhango District*

How to cite this work (APA):

Dushimimana, E. & Habiyaremye, J. D. (2026). Evaluating the Capacity of Biology Multiple-Choice Questions to Assess Higher-order Thinking Skills in Secondary Schools: A Case Study of Ruhango District, Rwanda. *Journal of Research Innovation and Implications in Education, 10*(3), 3111 – 3124. **https://doi.org/10.59765/bgf**

1. Introduction

Assessment shapes what and how students learn, guiding instructional decisions and measuring the extent to which curriculum goals are achieved(Abduvaliyeva & Gaziyeva, 2025) . Globally, education is shifting away from rote memorization toward cultivating Higher-Order Thinking Skills (HOTS)—the analyzing, evaluating, and creating processes of Bloom's Revised Taxonomy that enable learners to apply knowledge meaningfully(Thankachan, 2024). Although Multiple-Choice Questions (MCQs) are widely believed to be limited to recall and comprehension, emerging evidence shows that carefully designed MCQs incorporating real-life contexts, data interpretation, and problem-based scenarios can effectively assess HOTS (Haladyna & Rodriguez, 2021; Kumar et al., 2024).

Rwanda's adoption of the Competence-Based Curriculum (CBC) in 2015 marked a deliberate shift from content-driven, memorization-based instruction toward developing critical thinking, problem-solving, creativity, communication, and collaboration(REB, 2015). Biology, given its emphasis on interpreting data, analyzing biological systems, and solving real-life problems, offers a natural context for cultivating HOTS(Rachael et al., 2024). MCQs remain the dominant assessment format in Rwandan classrooms and national examinations because of their efficiency and objectivity (Dumbraveanu et al., 2025), yet it remains uncertain whether these items, particularly those used by the National Examination and School Inspection Authority (NESA), adequately assess HOTS rather than mere recall (Gjika et al., 2024).

National data show that the O-Level Biology pass rate for the 2024/2025 academic year stood at only 44.75%, raising concern about learners' demonstrated critical thinking and problem-solving abilities(MINEDUC& NESA, 2025). This concern is intensified in rural districts such as Ruhango, where overcrowded classrooms, limited laboratories, and variable teacher training constrain the design and use of cognitively demanding assessments(Mugiraneza, J. P., 2021). Limited empirical evidence exists on how Rwandan Biology teachers design and use MCQs to assess HOTS, leaving a critical gap in understanding whether current assessment practice aligns with CBC goals(Elgadal & Mariod, 2021; Nshimiyimana & Andala, 2024).

1.1 Statement of the Problem

Although Rwanda's Competence-Based Curriculum explicitly requires classroom and school-based assessment to measure learners' higher-order thinking, and although Biology is well suited to cultivating such skills through its emphasis on data interpretation and real-life problem-solving, it remains unclear whether the Biology MCQs actually administered in secondary schools fulfil this requirement. The persistently low O-Level Biology pass rate of 44.75% for the 2024/2025 academic year, together with the rural constraints characteristic of Ruhango District, including large class sizes, limited laboratory facilities, and variable teacher training, suggest that assessment practice may still default to rote recall rather than to the cognitively demanding items the CBC envisions. However, no study has empirically classified the cognitive levels of Biology MCQs used in Ruhango District against Bloom's Revised Taxonomy, nor examined, alongside this classification, the specific knowledge, perceptions, and institutional and resource constraints that shape teachers' capacity to construct HOTS-oriented items. In the absence of such evidence, curriculum implementers, school leaders, and teachers lack a concrete basis for judging whether current Biology assessment practice supports or undermines the CBC's competence-based goals. It is this problem that the present study addresses.

This study therefore evaluated the extent to which MCQs used in secondary school Biology assessments in Ruhango District measure HOTS and examined the factors influencing their design and effectiveness under the CBC. Specifically, the study sought to: (i) analyze the cognitive levels of Biology MCQs against Bloom's Revised Taxonomy; (ii) examine Biology teachers' knowledge, perceptions, and practices regarding HOTS-oriented MCQs; and (iii) identify the pedagogical, institutional, and resource factors affecting MCQs' capacity to assess HOTS. The findings are expected to inform teacher training, school-level assessment moderation, and national assessment policy, while contributing to the limited literature on aligning classroom assessment with competence-based curricula in developing-country contexts(Kwok, 2021).

2. Literature Review

MCQs are widely used in education because of their efficiency, objectivity, and suitability for large classes(Casu & García-García, 2018; Rodriguez & Haladyna, 2013). Whether they can adequately assess HOTS, particularly in subjects such as Biology that demand analysis and evaluation of evidence, remains debated (Malamed, 2021). In North America, MCQs feature prominently in high-stakes tests such as the SAT and AP examinations, which are designed to probe analytical as well as retentive abilities (McKenna, 2019). In parts of Asia, heavy reliance on MCQ-based entrance examinations has been criticized for encouraging rote memorization over deep understanding(Chen & Teo, 2020). Studies analyzing large samples of biology assessment items consistently find that most target recall or comprehension: (Crowe et al., 2008) found about 70% of biology exam items assessed recall or comprehension, a pattern corroborated by (Momsen et al., 2010).

In Rwanda, national assessment analyses reveal a similar pattern, with roughly 68% of MCQs emphasizing recall-based knowledge over evaluative skills, and a persistent gap in empirically validated HOTS-oriented items for rural classrooms (Ndihokubwayo & Habiyaremye, 2018). Reports from the National Examination and School Inspection Authority (2023) similarly show that many national examination items emphasize recall and routine procedure rather than complex reasoning. While well-constructed MCQs are internationally recognized as capable of measuring higher-order cognition(Scully, 2017), few studies empirically validate how such items

function in real classroom examinations or how reliably teachers construct them, particularly within Rwanda's CBC context.

The CBC represents a substantial shift toward hands-on, learner-centered, competency-focused instruction and assessment(Ministry of Education Rwanda, 2015; REB, 2015). Translating this policy into classroom practice, however, is constrained by infrastructural and resource deficits and limited continuous professional development, especially in rural schools(Nsengimana et al., 2023),(Byusa et al., 2020). Theoretically, Bloom's Revised Taxonomy(Anderson & Krathwohl, 2001) anchors the analysis of MCQs' cognitive demand, distinguishing recall-oriented Remembering, Understanding, and Applying levels (LOTS) from the Analyzing, Evaluating, and Creating levels (HOTS) that require learners to interpret evidence, judge validity, and generate original solutions (Zaidi et al., 2018). Constructivist theory further supports designing MCQs around authentic scenarios that require active problem-solving rather than passive recall (Wibowo et al., 2025).

Structurally, MCQs comprise a stem, a correct answer, and distractors; conventional items tend to be cognitively simple, but scenario-based or data-rich stems can substantially raise their cognitive demand (DiBattista & Kurzawa, 2011), (Kurdi et al., 2021). Effective HOTS-oriented items typically employ plausible, misconception-based distractors and reasoning-focused stems (Wise, 2020). Teachers' assessment literacy is central to this process: those with stronger understanding of Bloom's Taxonomy and assessment principles are more likely to design items that go beyond recall, whereas limited literacy tends to produce recall-dominated items (Driana & Ernawati, 2019). Pedagogical, institutional, and resource-related factors - including access to item banks, instructional materials, and continuous professional development - further shape teachers' capacity to design HOTS-oriented MCQs (Gupta et al., 2020).

Taken together, the reviewed literature converges on three points relevant to this study. First, MCQs are not inherently limited to LOTS; their cognitive demand is a function of deliberate item-writing design rather than of the format itself (Scully, 2017), (Wise, 2020). Second, teachers' assessment literacy and access to professional development and item-writing resources are consistently identified as the strongest determinants of whether MCQs actually achieve HOTS-level demand in practice (Driana & Ernawati, 2019), (Gupta et al., 2020). Third, while Rwanda's CBC formally mandates competence-based, HOTS-oriented assessment, most available evidence on its implementation addresses curriculum delivery rather than the cognitive quality of the assessment instruments teachers actually use in Biology classrooms (Nsengimana et al., 2023), (Moses & Saine, 2025). This study addresses that gap directly by combining document analysis of actual Biology MCQs with teachers' and school leaders' accounts of the factors shaping how those items are constructed.

3. Methodology

This study adopted a pragmatic research paradigm, which supports integrating quantitative and qualitative evidence to address complex, real-world educational problems (Elgeddawy & Abouraia, 2024; Foster, 2024). A convergent mixed-methods approach and a case study design were used to evaluate MCQs used in secondary school Biology assessments in Ruhango District, Southern Province, Rwanda- a predominantly rural district hosting a diverse mix of public, government-aided, and private day and boarding schools implementing the CBC.

Three complementary data sources were used. First, document analysis was conducted on archived Biology MCQs drawn from classroom tests and end-of-term examinations administered between 2024 and 2026. From an initial pool of approximately 300 eligible papers meeting inclusion criteria (officially administered, archived, legible), one paper per grade level (S1–S6) was randomly selected from each of twelve purposively sampled schools, yielding 681 MCQs for cognitive classification against Bloom's Revised Taxonomy (Anderson & Krathwohl, 2001). Second, structured questionnaires combining closed-ended and Likert-scale items were administered to 16 Biology teachers, selected through stratified purposive sampling to capture variation in school type, location, and experience. Third, semi-structured interviews were conducted with 17 school leaders (8 Directors of Studies and 9 Headteachers) across the sampled schools to explore institutional, pedagogical, and resource-related factors. The overall respondent sample size ($n \approx 45$) was determined using Yamane's (1967) simplified formula at a 95% confidence level and 5% margin of error.

Quantitative data were analyzed descriptively using frequencies, percentages, means, and standard deviations computed in JAMOVI and STATA. Qualitative data from interviews were analyzed thematically following Braun and Clarke's six-step framework (Naeem et al., 2023), encompassing familiarization, coding, theme development, review, definition, and reporting. Triangulating document analysis, teacher questionnaires, and school-leader interviews strengthened the credibility and trustworthiness of the findings(Donkoh, 2023).

A pilot study involving 16 participants with similar characteristics, but excluded from, the final sample assessed the instrument's clarity and reliability. All

subscales exceeded the 0.75 Cronbach's alpha threshold, ranging from α = 0.853 (Effectiveness of MCQs in Assessing HOTS) to α = 0.934 (Teachers' Perceptions of MCQs), indicating strong internal consistency. Ethical clearance was obtained from the University of Rwanda Ethics Committee, with authorization from MINEDUC, district education authorities, and school administrations. Participation was voluntary and based on written informed consent; data were anonymized, treated confidentially, and reported in aggregate form (UR, 2020; MINEDUC, 2021).

4. Results and Discussion

4.1 Cognitive Levels of Biology MCQs against Bloom's Revised Taxonomy

Table 1 presents the distribution of the 681 analyzed MCQs across the six levels of Bloom's Revised Taxonomy.

Table 1: Distribution of Biology MCQs Across Bloom's Revised Taxonomy Levels (N = 681)

Bloom's Level	Frequency	Percentage	Cognitive Category
Remembering	570	83.7%	LOTS: 672 (98.7%)
Understanding	95	14.0%	
Applying	7	1.0%	
Analyzing	8	1.2%	HOTS: 9 (1.3%)
Evaluating	1	0.1%	
Creating	0	0.0%	

Note. Document analysis, 2026.

Biology MCQs used in the sampled schools overwhelmingly targeted lower-order cognitive skills, with Remembering alone accounting for 83.7% of items and understanding a further 14.0%. Only 1.3% of items assessed HOTS, and no item required Creating. This pattern mirrors international evidence: (Crowe et al., 2008) and (Momsen et al., 2010) similarly found that most biology assessment items emphasize recall over analysis and evaluation, while Rwandan national data show comparable emphasis on recall-based items (Ndihokubwayo & Habiyaremye, 2018); (MINEDUC& NESA, 2025). The near-total absence of HOTS items suggests that current Biology assessment practice in Ruhango District falls markedly short of the CBC's expectation that assessments should measure critical thinking, evaluation, and creativity (REB, 2015).

School-level analysis (Table 2) confirmed this pattern across all twelve sampled schools, with LOTS proportions ranging from 96.4% to 100% and several schools recording no HOTS items at all.

Table 2: School-Level Distribution of LOTS and HOTS Biology MCQs

School	LOTS f (%)	HOTS f (%)	Total
School A	110 (99.1%)	1 (0.9%)	111
School B	147 (98.0%)	3 (2.0%)	150
School C	54 (96.4%)	2 (3.6%)	56
School D	98 (97.0%)	3 (3.0%)	101
School E	62 (100%)	0 (0.0%)	62
School F	48 (100%)	0 (0.0%)	48
School G	20 (100%)	0 (0.0%)	20
School H	32 (100%)	0 (0.0%)	32
School I	65 (100%)	0 (0.0%)	65
School J	36 (100%)	0 (0.0%)	36
Overall	672 (98.7%)	9 (1.3%)	681

Note. Document analysis, 2026.

Analysis by grade level showed that Senior One and Senior Two examinations were almost entirely recall-oriented, requiring learners to identify, define, or classify biological structures and organisms. Senior Three papers showed only modest improvement, remaining dominated by items on human physiology framed as identification or definition tasks despite the topics' potential for analytical questioning. Some improvement emerged in Senior Five and Senior Six papers, where a limited number of ecology and genetics items required learners to interpret trophic relationships or experimental data, reflecting genuine HOTS characteristics; nevertheless, such items remained a small minority. Only one item across the entire sample required learners to evaluate competing biological explanations, and none required hypothesis generation or experimental design-the Creating level of Bloom's Taxonomy.

This concentration of cognitive demand at the lower end in Senior One and Two, with only marginal improvement toward Senior Five and Six, suggests that HOTS-oriented item writing in the sampled schools is opportunistic rather than systematic: it appears where topics such as ecology or genetics lend themselves naturally to interpretation, rather than resulting from deliberate item-writing design applied consistently across the curriculum. This pattern is consistent with (Haladyna & Rodriguez, 2013) observation that item writers default to recall-based formats unless specifically trained to design higher-order alternatives, and it reinforces (Scully, 2017) caution that the theoretical capacity of MCQs to assess HOTS does not automatically translate into practice without deliberate item construction. The virtual absence of Evaluating and Creating items across all twelve schools indicates that Biology assessment in Ruhango District, as currently practiced, cannot be relied upon to certify the higher competencies that the CBC identifies as central learning outcomes, regardless of what curriculum documents themselves prescribe (REB, 2015).

4.2 Teachers' Knowledge, Perceptions, and Practices Regarding HOTS-Oriented MCQs

Most teacher respondents (93.8%) held a Bachelor's degree, with teaching experience ranging from under five years (43.8%) to more than fifteen years (25.0%), and the majority (62.5%) teaching at Lower Secondary level, where CBC emphasis on HOTS is strong.

Table 3: the teachers' self-reported knowledge, perceptions, and practices regarding the construction of Biology multiple-choice questions (MCQs) for assessing Higher-Order Thinking Skills (HOTS)

Teachers' knowledge, perceptions, and practices regarding MCQs for assessing	Percentage level of teachers' self-reported knowledge and practices regarding the construction of biology MCQs for assessing HOTS				
Statement	**Strongly agree (%)**	**Agree (%)**	**Neutral (%)**	**Disagree (%)**	**Strongly disagree (%)**
I clearly understand the differences between lower-order and higher-order cognitive levels when constructing MCQs.	12.5	62.5	18.8	0	6.3
I can confidently design MCQs that assess analysis, evaluation, and creation in Biology.	6.3	25.0	0	62.5	6.3
When designing MCQs, I intentionally plan items to target specific cognitive levels.	12.5	6.3	6.3	68.8	6.3
I use Bloom's Taxonomy (or similar frameworks) as a guide when constructing Biology MCQs.	12.5	0	18.8	56.3	12.5
Grand mean	10.95	23.45	10.975	46.9	7.85

Source: Primary data, 2026

The findings indicate that teachers possess a good theoretical understanding of higher-order cognitive skills but demonstrate limitations in applying this knowledge to practical assessment design. Specifically, 75.0% of the respondents agreed or strongly agreed that they clearly understand the differences between lower-order and higher-order cognitive levels, suggesting that most teachers are familiar with Bloom's Taxonomy and the

cognitive demands of the Competence-Based Curriculum. However, the grand mean percentages show that only 34.4% of teachers expressed overall agreement with the four statements, whereas approximately 54.8% expressed disagreement, indicating that practical implementation remains a challenge.

Despite demonstrating conceptual knowledge, teachers reported lower confidence in constructing MCQs that effectively assess higher-order thinking. Only 31.3% agreed that they could confidently design biology MCQs measuring analysis, evaluation, and creation, while 68.8% disagreed or strongly disagreed. Similarly, only 18.8% indicated that they intentionally plan MCQs to target specific cognitive levels, and the majority also reported limited use of Bloom's Taxonomy when constructing assessment items. These findings suggest that although teachers recognize the importance of HOTS-oriented assessment, many of them experience difficulties in translating theoretical knowledge into effective classroom assessment practices.

The findings imply that while biology teachers possess the foundational knowledge required for competency-based assessment, additional capacity-building initiatives are necessary to strengthen their practical assessment skills. The relatively low confidence in designing higher-order MCQs suggests that constructing valid and reliable assessment items requires specialized expertise, continuous practice, and sustained professional support.

Teachers' practices and constraints in using MCQs to assess HOTS

About the teachers' practices and constraints associated with the use of MCQs to assess HOTS within the framework of the Competence-Based Curriculum, the results revealed variations in challenges perceived by the teachers table 6. The contacted teachers were required to indicate the approaches they use during the time when they are designing MCQs aimed at assessing students' HOTS. As respondents were allowed to select more than one option, the percentages represent the proportion of teachers who reported using each approach. The findings are presented in table 6.

Table 4: Approaches used by teachers to design MCQs for assessing HOTS

S/N	Approach used when designing HOTS-oriented MCQs	Frequency	Percentage (%)
1	Use complex scenarios or case studies in the stem	**12**	**75.0**
2	Include plausible distractors that require evaluation	**12**	**75.0**
3	Require judgment based on criteria or evidence	**12**	**75.0**
4	Design MCQs that can assess creation through selecting the best hypothesis or experimental design	**8**	**50.0**
5	MCQs which include well-designed distractors that are scientifically plausible, reflect common misconceptions, are similar in length and structure to the correct answer, and are not obviously incorrect.	**11**	**68.8**
6	MCQs constructed to prevent guessing through test-wise strategies by avoiding absolute terms, overly long correct answers compared to distractors, and the repetition of stem wording only in the correct option.	**12**	**75.0**
8	Use HOT verbs HOTS verbs include analyze, compare, differentiate, interpret, Evaluate, justify, critique, predict Design, formulate, propose, conclude	**13**	**81.3**
9	Use stimulus-based questions that present new contextual information such as experimental data, graphs, case studies, or short descriptions of biological phenomena to assess students' ability to analyze and apply knowledge	**10**	**62.5**
10	Emphasize problem-solving	**12**	**75.0**
11	Link biology to real-life contexts	**16**	**100**
12	Encourage reasoning and decision-making	**13**	**81.3**
13	Use Bloom's Taxanomy checklist	**14**	**87.5**

Source: Primary data, 2026

The results revealed that the most frequently reported practice was linking biology concepts to real-life contexts, which was selected by 100% of the respondents. This was followed by the use of a Bloom's Taxonomy checklist (87.5%), while 81.3% reported using HOTS action verbs such as *analyzing, comparing, evaluating, justifying,* and *predicting*, indicating that they intentionally encourage reasoning and decision-making when constructing MCQs. These findings suggest that biology teachers are highly aware of the fundamental principles underpinning competency-based assessment and recognize the importance of contextualizing assessment tasks to promote meaningful learning. By connecting biology concepts to authentic situations and aligning questions with higher cognitive processes, teachers demonstrate an intention to develop learners' analytical reasoning, decision-making, and problem-solving abilities rather than merely assessing factual recall.

Several other HOTS-oriented practices were also widely reported. Approximately 75.0% of the respondents indicated that they use complex scenarios or case studies, include plausible distractors requiring evaluation, require students to make judgments based on evidence or established criteria, construct MCQs that minimize guessing through appropriate item-writing techniques, and emphasize problem-solving during question development. These findings indicate that a substantial proportion of biology teachers understand key principles of effective MCQ construction and attempt to design assessment items that challenge learners to analyze information, evaluate alternatives, and apply biological knowledge in unfamiliar situations. Such practices are consistent with the objectives of the Competence-Based Curriculum, which emphasizes inquiry, reasoning, and application of knowledge in solving real-world problems.

Despite these encouraging findings, some advanced HOTS-oriented assessment practices were implemented less consistently. Although 68.8% of the respondents reported constructing scientifically plausible distractors that reflect common misconceptions, only 62.5% indicated that they regularly use stimulus-based questions incorporating experimental data, graphs, case studies, or biological phenomena. Similarly, 62.5% reported applying advanced item-writing techniques that minimize test-wise guessing strategies. Most notably, only 50.0% of the respondents indicated that they design MCQs capable of assessing the highest cognitive level of creation, such as selecting the most appropriate hypothesis or experimental design. This was the least frequently reported practice, suggesting that teachers encounter greater difficulty in developing assessment items targeting the "Create" level of Bloom's Revised Taxonomy. Designing such questions requires high assessment literacy, deeper content knowledge, and greater expertise in constructing authentic scientific inquiry tasks than those required for assessing analysis or evaluation.

Furthermore, the qualitative results showed the use of numerous reflective strategies to assess if their MCQ effectively evaluate HOTS. The main reported strategies were evaluating questions against revised Bloom's taxonomy, asking learners to justify their answers with the purpose of examining students' reasoning, and analyzing students' answers to improve the quality of future questions. This is because teachers reported that designing HOTS-oriented MCQs requires continuous improvement.

Further biology teachers' responses regarding their practices, perceptions, and challenges about constructing and using HOTS-oriented MCQs in biology assessments under the CBC were given through a semi-structured interview. These particularly concern the approaches commonly used by the contacted biology teachers when designing HOTS-oriented MCQs and factors that negatively impact the development and use of such questions.

According to responses received from participants to this study, only a few teachers incorporate strategies like real-life biological cases, data interpretation items, experimental contexts to be interpreted, and problem-solving situations to be analyzed. Nearly one-third reported rarely using these approaches. The major barriers reported included large classes, inadequate training in HOTS-oriented assessment, insufficient sample HOTS-oriented MCQs to be used for model, and time constraints. Overall, the current study findings indicate that biology assessment practices is still highly focusing on lower-order cognitive skills (LOTS), underlining the need for strengthened teachers' capacity building and assessment reforms to improve the development and assessment of higher-order thinking skills (HOTS).

4.3 Pedagogical, Institutional, and Resource Factors Affecting HOTS Assessment

Thematic analysis of interviews with the 16 school leaders (7 DOS, 9 Headteachers) generated three themes (Table 5).

Table 5: Themes and Sub-Themes from School Leader Interviews

Theme	Sub-themes
Pedagogical factors	Assessment literacy; limited teaching experience; inadequate continuous professional development
Institutional factors	Leadership commitment; assessment moderation and quality assurance; curriculum implementation; assessment culture
Resource factors	Assessment materials and reference resources; time constraints; professional collaboration and mentoring

Note. Interview data, 2026.

Under pedagogical factors, respondents consistently linked teachers' assessment literacy and classroom experience to their capacity to design HOTS-oriented items. As one school leader explained, newly recruited teachers tend to rely heavily on recall questions because of limited assessment experience, while continuous professional development sessions were described as excessively theoretical rather than practice-based, with one respondent noting that teachers need trainings where they can practically construct, review, and revise MCQs collaboratively.

Under institutional factors, respondents emphasized the role of instructional leadership and examination moderation- departmental review against curriculum blueprints and NESA criteria-in sustaining assessment quality, while noting substantial variation in CBC implementation across schools linked to differences in leadership commitment. One respondent captured this by stressing that assessment should be viewed as a tool for improving learning rather than simply assigning grades.

Under resource factors, respondents identified limited access to validated item banks, assessment manuals, and reference materials, alongside insufficient time for collaborative item development given heavy teaching and administrative workloads, as key barriers; they recommended school-based and national HOTS item banks and stronger professional learning communities to support item development and moderation.

These qualitative findings converge with the quantitative results: the near-absence of HOTS items in the document analysis (Table 1) is explained by teachers' limited practical assessment competence (Section 4.2), which is itself shaped by inadequate professional development, weak institutional moderation systems, and scarce assessment resources and time consistent with prior findings that HOTS assessment quality depends on the interaction of teacher competence, institutional support, and resource availability (Setiyowati et al., 2025), (Nsengimana et al., 2023).

Discussion of the Results

This part discusses the study findings on "Evaluating biology Multiple Choice Questions' capacity to assess Higher-Order Thinking Skills in secondary schools' biology: A case study of Ruhango district, Rwanda". The study intended to (1) analyze the cognitive levels of biology MCQs in Ruhango secondary schools regarding higher-order thinking skills as outline in Bloom's revised taxonomy; (2) examine teachers' knowledge, perceptions, and practices on using MCQs to assess higher-order thinking under the CBC; and (3) identify pedagogical, institutional, and resource factors affecting MCQs' capacity in assessing higher-order thinking in biology. The discussion elements are organized according to these study specific objectives.

4.4 Cognitive levels of biology Multiple Choice Questions used in secondary schools in Ruhango district

Biology MCQs used in the sampled secondary schools of Ruhango district overwhelmingly assessed lower-order cognitive skills (LOTS). The assessment of the multiple-choice questions (MCQ) in papers administered across all sampled schools showed a predominant emphasis on LOTS. Across all the sampled schools, biology MCQs predominantly assessed LOTS, especially remembering and understanding and very few MCQs evaluate HOTS, and in some schools, no HOTS items were identified. Only a very small fraction of the questions given was formulated to assess Higher-Order Thinking Skills (HOTS).

This finding revealed that assessments in biology subject still emphasize just recall and basic comprehension rather than HOTS that include analysis, evaluation, and creation as stipulated by revised Bloom's taxonomy and the Competence-Based Curriculum (CBC). In this regard, students are not given opportunities to foster their critical thinking, scientific reasoning and problem-solving skills. These findings suggest that biology fall short of promoting the HOTS expected be gained through in competency-based education.

The present findings agree with the previous findings which report the predominance of lower-order cognitive questions in science assessments. (Bartolome & Navalta, 2025)Bartolome , (Goswami et al., 2025) and (Koman et al., 2023), similarly reported that most science examination questions and tests measure remembering and understanding, with limited assessment of HOTS. Similarly, (Kiran Riaz et al., 2025; Milla et al., 2025; Nadia Amari et al., 2025 and Prihastuti et al., 2020) attributed this pattern to limited teacher assessment literacy, reliance on traditional assessment practices, examination-oriented teaching, and the relative ease of constructing and scoring recall-based questions.

On the other hand,(Hikmah et al., 2023; Reinier M. Navarro, 2024; Solikah et al., 2025; and Zaiful Shah & Zakaria, 2024) demonstrated that planned professional development in assessment design, collaborative examination moderation, and the use of authentic, scenario-based assessment tasks substantially increase the proportion of HOTS-oriented MCQs and improve learners' critical thinking and scientific reasoning. Thus, the need for continuous professional development (CPD) to equip teachers with enough assessment literacy, systematic examination moderation, and deliberate integration of authentic biological contexts into MCQs.

Other educational scholars have emphasized that increasing the proportion of HOTS questions enhances students' critical thinking, conceptual understanding, scientific literacy, and academic achievement by encouraging learners to interpret evidence, solve unfamiliar problems, justify decisions, and synthesize knowledge from multiple sources. In a similar orientation, Solikah et al. (2025) demonstrated that schools providing continuous professional development on assessment design significantly increased the proportion of HOTS questions, resulting in improved analytical thinking, scientific reasoning, and problem-solving abilities among learners. These studies emphasize on the importance of equipping teachers with assessment competencies that support the objectives of competency-based education. Once done, it will ensure that biology assessments and tests align with the goals of the CBC by strengthening HOTS, scientific inquiry, and meaningful learning.

4.5 Biology teachers' Knowledge, perceptions, and practices on HOTS-Oriented Multiple-Choice Questions

The findings revealed that biology MCQs administered in Ruhango district secondary schools predominantly evaluated LOTS. These include particularly recalling and understanding, with limited focus on HOTS which analysis, evaluation, and creation. This is inconsistent with the predictions by the CBC advocating assessments that promote critical thinking and problem-solving. Our results agree with reported findings by (Liu et al., 2024). These reported a focus on factual recall by majority of biology assessments just because these questions are mainly easier to construct. On the other side, (Reinier M. Navarro, 2024) reported that appropriate MCQs must include real life scenarios and problem-based contexts to effectively assess HOTS. This underlines quality of the constructed MCQs in addition to the format itself.

The current study again showed that even if the contacted biology teachers possess good knowledge and positive perceptions about HOTS-oriented MCQs, their practical skills in designing the appropriate questions are still limited. These results concur with (Milla et al., 2025).These revealed that even though the teachers understand competency-based assessment, they require continuous professional development to foster their assessment literacy and item-writing skills. Similarly, (Liu et al., 2024) recommended continuous structured assessment training sessions to significantly strengthen teachers' capacity to develop HOTS-oriented MCQS. In the same way,(Setiyowati et al., 2025) highlighted that teacher competence, support provided by the institution, and continuous professional development are highly recommended for good HOTS assessment implementation.

Finally, the contacted teachers within the current study reported several challenges which interfered with developing HOTS-oriented MCQS. The challenges revealed include large class sizes, lack of training on assessment, inadequate preparation time, and limited instructional resources. These findings align with (Milla et al., 2025) ,who revealed limited teacher competence, insufficient institutional support, and inadequate available resources to negatively impact the successful implementation of HOTS-oriented assessment. Furthermore, although teachers generally recognize the value of MCQs for assessing higher-order thinking, (Hull & Vígh, 2025) highlighted that positive perceptions alone are not enough to comply with HOTS-oriented assessment and must be supported by continuous professional development.

4.6 Pedagogical, institutional, and resource factors affecting the use of MCQs' to assess Higher-Order Thinking in biology

The findings on pedagogical, institutional, and resource factors affecting the application of multiple-choice questions (MCQs) to assess Higher-Order Thinking Skills (HOTS) in biology indicate that teachers' assessment literacy, teaching experience, and practical professional development are fundamental to constructing biology

MCQs that assess higher-order thinking skills (HOTS). Participants emphasized that teachers with a sound understanding of the revised Bloom taxonomy and the Competence-Based Curriculum (CBC) are more capable of designing questions that promote analysis, evaluation, and problem-solving. However, existing Continuous Professional Development (CPD) programs were perceived as overly theoretical, limiting teachers' practical assessment skills. The results agree with past studies that have highlighted the importance of assessment literacy and hands-on professional training in the development of competency-based assessments(Brookhart, 2011).

The study further revealed that effective instructional leadership, examination moderation, and a positive assessment culture strengthen the quality of HOTS-oriented MCQs. School leaders were reported to play a significant role in promoting curriculum implementation, quality assurance, and adherence to national assessment standards. This finding supports research indicating that strong instructional leadership and moderation improve assessment quality and curriculum implementation(Gatela & Arpilleda, 2025).

Finally, inadequate assessment resources, limited time, and weak professional collaboration were identified as major barriers to developing quality MCQs. Participants recommended establishing validated item banks, increasing collaboration among teachers, and providing adequate time for assessment design and moderation. The findings are consistent with research emphasizing the value of resources, collaborative learning, and institutional support for effective assessment (DeLuca et al., 2023).

5. Conclusion and Recommendations

5.1 Conclusion

This study concludes that Biology MCQs used in secondary schools in Ruhango District remain overwhelmingly knowledge-based, with 98.7% of analyzed items assessing Lower-Order Thinking Skills and only 1.3% assessing Higher-Order Thinking Skills, falling short of the Competence-Based Curriculum's expectation of balanced cognitive assessment (Anderson & Krathwohl, 2001). Although Biology teachers demonstrated reasonable theoretical understanding of HOTS and generally positive perceptions of MCQs' potential to assess higher-order cognition, they lacked the practical assessment competence and institutional support needed to translate this understanding into classroom practice, constrained by inadequate professional development, large classes, limited resources, and time pressure(Brookhart, 2011).

5.2 Recommendations

Based on these findings, the study recommends that:

1. Teacher training institutions strengthen pre-service and in-service training in assessment literacy, with practical emphasis on constructing scenario-based, Bloom-aligned MCQs;
2. MINEDUC, REB, and NESA develop continuous professional development programmes and require balanced cognitive-level distributions in national and school-based examinations;
3. Headteachers and Directors of Studies institute regular assessment moderation using Bloom's Revised Taxonomy prior to examination administration;
4. Biology teachers deliberately increase HOTS-oriented items by incorporating authentic scenarios, data interpretation, and problem-solving tasks, and collaborate through professional learning communities;
5. Schools provide adequate instructional and assessment resources and allocate time for collaborative test construction;
6. Curriculum developers produce exemplar HOTS-aligned Biology MCQ banks; and
7. Future researchers examine the relationship between HOTS-oriented MCQs and students' academic achievement across broader samples and districts.

References

Abduvaliyeva, O., & Gaziyeva, S. (2025). *ASSESSMENT: CONCEPT, TYPES, AND MODERN APPROACHES*. https://doi.org/10.5281/ZENODO.15550632

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

Brookhart, S. M. (2011). Educational Assessment Knowledge and Skills for Teachers: Spring 2011. *Educational Measurement: Issues and Practice*, *30*(1), 3–12. https://doi.org/10.1111/j.1745-3992.2010.00195.x

Byusa, E., Kampire, E., & Mwesigye, A. R. (2020). Analysis of Teaching Techniques and Scheme of Work in Teaching Chemistry in Rwandan Secondary Schools. *Eurasia Journal of Mathematics, Science and Technology Education*, *16*(6). https://doi.org/10.29333/ejmste/7833

Casu, G., & García-García, C. (2018). Differential Length and Overlap with the Stem in Multiple-Choice Item Options: A Pilot Experiment. *Psicología Educativa*, *25*(1), 43–48. https://doi.org/10.5093/psed2018a20

Chen, J., & Teo, T. (2020). Chinese school teachers' conceptions of high-stakes and low-stakes assessments: An invariance analysis. *Educational Studies*, *46*(4), 458–475. https://doi.org/10.1080/03055698.2019.1599823

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

DiBattista, D., & Kurzawa, L. (2011). Examination of the Quality of Multiple-choice Items on Classroom Tests. *The Canadian Journal for the Scholarship of Teaching and Learning*, *2*(2). https://doi.org/10.5206/cjsotl-rcacea.2011.2.4

Donkoh, S. (2023). Application of triangulation in qualitative research. *Journal of Applied Biotechnology & Bioengineering*, *10*(1), 6–9. https://doi.org/10.15406/jabb.2023.10.00319

Driana, E., & Ernawati, E. (2019). TEACHERS' UNDERSTANDING AND PRACTICES IN ASSESSING HIGHER ORDER THINKING SKILLS AT PRIMARY SCHOOLS. *ACITYA Journal of Teaching & Education*, *1*(2), 110–118. https://doi.org/10.30650/ajte.v1i2.233

Dumbraveanu, R., Baciu, V., Grosseck, G., & Milencovici, D. (2025). Multiple-Choice Tests: Objectivity or Delusion in Assessment? A Comparative Analysis from Romania and Moldova. *Revista Romaneasca Pentru Educatie Multidimensionala*, *17*(3), 505–531. https://doi.org/10.18662/rrem/17.3/1032

Elgadal, A. H., & Mariod, A. A. (2021). Item Analysis of Multiple-choice Questions (MCQs): Assessment Tool For Quality Assurance Measures. *Sudan Journal of Medical Sciences*, *16*(3), 334–346. https://www.ajol.info/index.php/sjms/article/view/216888

Elgeddawy, M., & Abouraia, M. (2024). Pragmatism as a Research Paradigm. *European Conference on Research Methodology for Business and Management Studies*, *23*(1), 71–74. https://doi.org/10.34190/ecrm.23.1.2444

Foster, C. (2024). Methodological pragmatism in educational research: From qualitative-quantitative to exploratory-confirmatory distinctions. *International Journal of Research & Method in Education*, *47*(1), 2. https://doi.org/10.1080/1743727X.2023.2210063

Gjika, E., Basha, L., & Alizoti, A. (2024). *EXPLORING THE DYNAMICS OF ASSESSMENTS: A COMPREHENSIVE REVIEW OF MCQs, SAQs, AND E- ASSESSMENT IN ALBANIA*. 2.

Gupta, P., Meena, P., Khan, A., Malhotra, R., & Singh, T. (2020). Effect of faculty training on quality of multiple-choice questions. *International Journal of Applied and Basic Medical Research*, *10*(3), 210. https://doi.org/10.4103/ijabmr.IJABMR_30_20

Haladyna, T. M., & Rodriguez, M. C. (2013). *Developing and Validating Test Items* (0 ed.). Routledge. https://doi.org/10.4324/9780203850381

Haladyna, T. M., & Rodriguez, M. C. (2021). Using Full-information Item Analysis to Improve Item Quality. *Educational Assessment*, *26*(3), 198–211. https://doi.org/10.1080/10627197.2021.1946390

Hull, P., & Vígh, T. (2025). Teachers' assessment literacy: A descriptive literature review. *Hungarian Educational Research Journal*, *15*(2), 210–225. https://doi.org/10.1556/063.2024.00317

Kumar, A. P., Nayak, A., K, M. S., Chaitanya, & Ghosh, K. (2024). A Novel Framework for the Generation of Multiple Choice Question Stems Using Semantic and Machine-Learning Techniques. *International Journal of Artificial Intelligence in Education*, *34*(2), 332–375. https://doi.org/10.1007/s40593-023-00333-6

Kurdi, G., Leo, J., Matentzoglu, N., Parsia, B., Sattler, U., Forge, S., Donato, G., & Dowling, W. (2021). A comparative study of methods for a priori prediction of MCQ difficulty. *Semantic Web*, *12*(3), 449–465. https://doi.org/10.3233/SW-200390

Kwok, P. K. P. (2021). *The implementation of learner-centred pedagogy in Rwanda: Teachers as mediators* [Apollo - University of Cambridge Repository]. https://doi.org/10.17863/CAM.83227

Malamed, C. (2021, November 4). Writing Multiple Choice Questions For Higher Order Thinking. *The eLearning Coach*. https://theelearningcoach.com/elearning_design/higher-order-multiple-choice-questions/

McKenna, P. (2019). Multiple choice questions: Answering correctly and knowing the answer. *Interactive Technology and Smart Education*, *16*(1), 59–73. https://doi.org/10.1108/ITSE-09-2018-0071

MINEDUC& NESA. (2025). *MINEDUC Publishes 2024/2025 National Examination Results for PLE and S3*. https://www.nesa.gov.rw/online-services/mineduc-publishes-2024-2025-national-examination-results-for-ple-and-s3

Ministry of Education Rwanda. (2015). *National Education for All 2015 Review; 2015*.

Momsen, J. L., Long, T. M., Wyse, S. A., & Ebert-May, D. (2010). Just the Facts? Introductory Undergraduate Biology Courses Focus on Low-Level Cognitive Skills. *CBE—Life Sciences Education*, *9*(4), 435–440. https://doi.org/10.1187/cbe.10-01-0001

Moses, M., & Saine, E. (2025). Competence Based Curriculum Implementation in Rwanda: Strategies and Challenges. *International Journal of Research and Innovation in Social Science*, *9*(11), 7189–7202. https://doi.org/10.47772/IJRISS.2025.91100558

Mugiraneza, J. P. (2021). *Digitalization in teaching and education in Rwanda: Background report.*

Naeem, M., Ozuem, W., Howell, K., & Ranfagni, S. (2023). A Step-by-Step Process of Thematic Analysis to Develop a Conceptual Model in Qualitative Research. *International Journal of Qualitative Methods*, *22*, 16094069231205789. https://doi.org/10.1177/16094069231205789

Ndihokubwayo, K., & Habiyaremye, H. T. (2018). Why did Rwanda shift from knowledge to competence based curriculum? Syllabuses and textbooks point of view. *African Research Review*, *12*(3), 38. https://doi.org/10.4314/afrrev.v12i3.4

Nsengimana, T., Mugabo, L. R., Ozawa, H., & Nkundabakura, P. (2023). Science Teachers' Knowledge, Understanding and Perceptions of Competence-Based Curriculum in Three Secondary Schools in Rwanda. *European Journal of Educational Research*, *volume–12–2023*(volume–12–issue–1–january–2023), 317–327. https://doi.org/10.12973/eu-jer.12.1.317

Nshimiyimana, V., & Andala, H. O. (2024). Effect of Implementing a Competence-Based Curriculum on Students' Academic Performance in Mathematics: A Case of Selected Lower Secondary Public Day Schools in Rwamagana District, Rwanda. *African Journal of Empirical Research*, *5*(4), 512–519. https://doi.org/10.51867/ajernet.5.4.41

Rachael, O.-J., Sanusi, D. B. O., & Omowale, and A. T. (2024). Post-Performance Assessment of Nigerian Centre for Disease Control (NCDC) Communication Strategies for Adherence to COVID-19 Safety Protocols among Non-Literates in Southwest Nigeria. *Journal of Mass Communication Department, Dept of Mass Communication, University of Karachi*, *31*. http://www.jmcd-uok.com/index.php/jmcd/article/view/389

REB. (2015). *SIMPLIFIED GUIDE TO CLASSROOM-BASED AND NATIONAL ASSESSMENTS IN RWANDA*.

Rodriguez, M. C., & Haladyna, T. M. (2013). Objective testing of educational achievement. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J.-I. C. Hansen, N. R. Kuncel, S. P. Reise, & M. C. Rodriguez (Eds.), *APA handbook of testing and assessment in psychology, Vol. 1: Test theory and testing and assessment in industrial and organizational psychology.* (pp. 305–314). American Psychological Association. https://doi.org/10.1037/14047-018

Scully, D. (2017). Constructing multiple-choice items to measure higher-order thinking. *Practical Assessment, Research and Evaluation (PARE)*, *22*(1), 1–13. https://doi.org/10.7275/swgt-rj52

Setiyowati, E., Puspita, A. M. I., & Mariana, N. (2025). Teachers' Pedagogical Competence in Designing Deep Learning and HOTS Assessment in Elementary Schools: A Systematic Literature Review. *Journal of Innovation and Research in Primary Education*, *4*(4), 4011–4029. https://doi.org/10.56916/jirpe.v4i4.2633

Thankachan, K. (2024). Paradigm shift from rote learning to critical thinking, experiential learning, and holistic development in the indian education

system. *Journal of Management Research and Analysis*, *11*, 140–141. https://doi.org/10.18231/j.jmra.2024.023

Wibowo, S., Wangid, M. N., & Firdaus, F. M. (2025). The relevance of Vygotsky's constructivism learning theory with the differentiated learning primary schools. *Journal of Education and Learning (EduLearn)*, *19*(1), 431–440. https://doi.org/10.11591/edulearn.v19i1.21197

Wise, M. J. (2020). An Empirical Investigation of the Fairness of Multiple-Choice Items Relative to Constructed-Response Items on Tests of Students' Mastery of Course Content. *Journal of Education, Teaching and Social Studies*, *2*(4), p16. https://doi.org/10.22158/jetss.v2n4p16

Zaidi, N. L. B., Grob, K. L., Monrad, S. M., Kurtz, J. B., Tai, A., Ahmed, A. Z., Gruppen, L. D., & Santen, S. A. (2018). Pushing Critical Thinking Skills With Multiple-Choice Questions: Does Bloom's Taxonomy Work? *Academic Medicine*, *93*(6), 856–859. https://doi.org/10.1097/ACM.0000000000002087