Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Associations between voluntary engagement with faculty-developed, clinically oriented formative questions and content-matched examination performance in preclinical microbiology: an observational mixed-methods cohort study

Hari Prasad Nepal

hnepal@burrell.edu

Burrell College School of Osteopathic Medicine

Research Article

Keywords: Formative assessment, Practice questions, Student engagement, Medical education, Microbiology education, Examination performance

Posted Date: November 16th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11260168/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Associations between voluntary engagement with faculty-developed, clinically oriented formative questions and content-matched examination performance in preclinical microbiology: an observational mixed-methods cohort study

Hari Prasad Nepal[1]

[1]Department of Biomedical Sciences, Burrell College of Osteopathic Medicine, Melbourne, FL, USA

Corresponding author:

Hari Prasad Nepal

Department of Biomedical Sciences, Burrell College of Osteopathic Medicine, Melbourne, FL, USA

Email: hnepal@burrell.edu

Abstract

Background

Formative practice questions may support retrieval, knowledge application, and self-regulated learning. However, there is limited evidence regarding voluntary engagement with faculty-developed, clinically oriented questions aligned with course assessments. This study aimed to examine the associations between engagement with faculty-developed formative microbiology questions and content-matched examination performance among first-year osteopathic medical students and explored students' perceptions of these resources.

Methods

This quantitative-priority, observational mixed-methods cohort study was conducted at a US osteopathic medical school during the 2025–2026 academic year. The quantitative analyses included 314 students with complete engagement and examination data. Students voluntarily accessed 174 faculty-developed, clinically oriented multiple-choice questions across 11 microbiology topics through InteDashboard, with explanatory feedback provided after each attempt. Engagement was measured as the percentage of questions completed, and the primary outcome was performance on corresponding summative examination items. Associations were primarily examined via Pearson correlation, simple linear regression, and analysis of variance. Deidentified responses to two instructor evaluation questions were thematically analyzed and integrated with quantitative findings.

Results

The students completed a mean of 128.3 ± 47.6 questions (73.7 ± 27.4%), and the mean examination performance was 79.4 ± 11.1%. Engagement was positively associated with examination performance ($r = 0.448$, $p < 0.001$), accounting for 20.1% of the observed variance

(R^2 = 0.201). Each additional question completed was associated with 0.037 additional correct examination items (95% CI: 0.0285–0.0448; $p < 0.001$). Mean performance increased across the low-, moderate-, and high-engagement groups (71.8%, 78.0%, and 84.0%, respectively; $p < 0.001$). Positive associations were observed across all 11 topics after correction for multiple comparisons. Thematic analysis of 72 eligible anonymous response records identified four themes: assessment readiness and curricular alignment; reinforcement and application of knowledge; metacognitive calibration and explanatory feedback; and conditions affecting educational value.

Conclusions

Greater voluntary engagement with faculty-developed, clinically oriented formative microbiology questions was associated with higher performance on content-matched summative examination items. The students' perceptions provided complementary insights into the resource's educational value. Although causality cannot be inferred from this observational study, the findings support the potential educational value of thoughtfully designed, curriculum-aligned formative-question resources in preclinical medical education.

Keywords

Formative assessment, Practice questions, Student engagement, Medical education, Microbiology education, Examination performance

Introduction

Formative assessment is an integral component of medical education, providing learners with opportunities to evaluate their understanding, identify knowledge gaps, and guide subsequent study. When aligned with learning objectives and accompanied by actionable feedback, formative assessment can support learning while helping learners monitor progress toward intended educational outcomes [1]. Retrieval practice is one mechanism through which practice questions may facilitate learning, as actively retrieving information can improve subsequent retention compared with restudying the same material [2]. Similarly, evidence from a systematic review of health profession education suggests that retrieval and distributed practice are associated with improved academic outcomes, although the interventions, assessment approaches, and study designs varied across the included studies [3].

In addition to knowledge retrieval, formative questions may support self-regulated learning by prompting students to monitor their understanding, recognize uncertainty, and identify areas requiring further study. A meta-analysis of undergraduate and graduate medical education reported a positive association between self-regulated learning strategies and learning outcomes, although the magnitude of this relationship varied across educational contexts and outcomes [4]. The potential educational value of formative assessment also depends on question design, alignment with curricular objectives, and feedback quality. Explanatory feedback may be especially useful because rationales can help learners understand why an answer is correct and apply that understanding to new questions [5]. Consistent with this perspective, a study of preclinical medical

students revealed that post examination feedback containing question rationales was associated with greater improvement in retesting than was feedback that provided the correct response alone [6].

Clinically oriented multiple-choice questions can be designed to assess the application and integration of foundational knowledge within clinical contexts, although the cognitive processes elicited depend on item construction and the demands of the question [7]. The development of clinical reasoning during preclinical medical education is also an established educational priority [8]. This emphasis on knowledge application is consistent with the COMLEX-USA Blueprint, which includes assessment of foundational knowledge and its application to osteopathic medical practice and clinical presentations [9]. Accordingly, clinically oriented formative questions may provide opportunities for learners to apply foundational knowledge in ways that align with broader application-oriented expectations in osteopathic medical education.

In microbiology, clinically contextualized questions may require students to connect microbial characteristics and mechanisms of pathogenesis with host responses, diagnostic findings, and clinical manifestations. Educational guidance has emphasized active learning, conceptual understanding, and clinical application in preclinical microbiology and infectious diseases [10]. Nevertheless, a national survey of U.S. medical schools identified challenges in the organization of microbiology curricula and the integration of foundational and clinical content [11]. Faculty-developed formative questions aligned with course learning objectives and accompanied by explanatory feedback may therefore provide structured opportunities for students to retrieve and apply microbiological knowledge while identifying areas for additional study.

Research examining online formative assessments and question banks has reported associations between student engagement and summative academic performance, although the evidence differs according to resource type, educational setting, and study design. A systematic review and meta-analysis of peer-generated multiple-choice question banks revealed a positive association between question-bank use and summative examination performance while recognizing limitations of the underlying observational evidence and the potential influence of confounding factors [12]. Individual studies of student-authored and cocreated question banks, including PeerWise, have also reported associations between engagement and examination outcomes, alongside the perceived educational benefits of question answering, question authoring, and peer discussion [13, 14]. However, findings from peer-generated resources may not directly generalize to faculty-developed questions that are explicitly aligned with course learning objectives, are clinically contextualized, and are accompanied by explanatory feedback. The relationship between voluntary engagement with such faculty-developed resources and performance on corresponding summative examination items remains less well characterized, particularly in preclinical microbiology education.

The present study therefore examined the associations between voluntary engagement with faculty-developed, clinically oriented formative practice questions and performance on content-

matched summative examination items among first-year osteopathic medical students in a preclinical microbiology curriculum. It also explored topic-specific associations and students' written perceptions of the resource's educational value. The study hypothesized that greater engagement would be positively associated with performance on corresponding summative examination items. Given the observational design and voluntary nature of resource use, the findings were interpreted as associations rather than evidence of causality.

Methods

Study Design and Setting

A quantitative-priority, observational mixed-methods cohort study was conducted among first-year osteopathic medical students at Burrell College School of Osteopathic Medicine during the 2025–2026 academic year. The quantitative component examined the associations between voluntary engagement with faculty-developed, clinically oriented formative practice questions and performance on content-matched summative examination items. On the other hand, the qualitative component analyzed deidentified responses to two open-ended questions from a routine instructor evaluation. The components were analyzed separately and integrated during interpretation. Because use of the formative resource was voluntary and no intervention was assigned, the quantitative findings were interpreted as associations rather than causal effects. The study design and analytic workflow are presented in Figure 1.

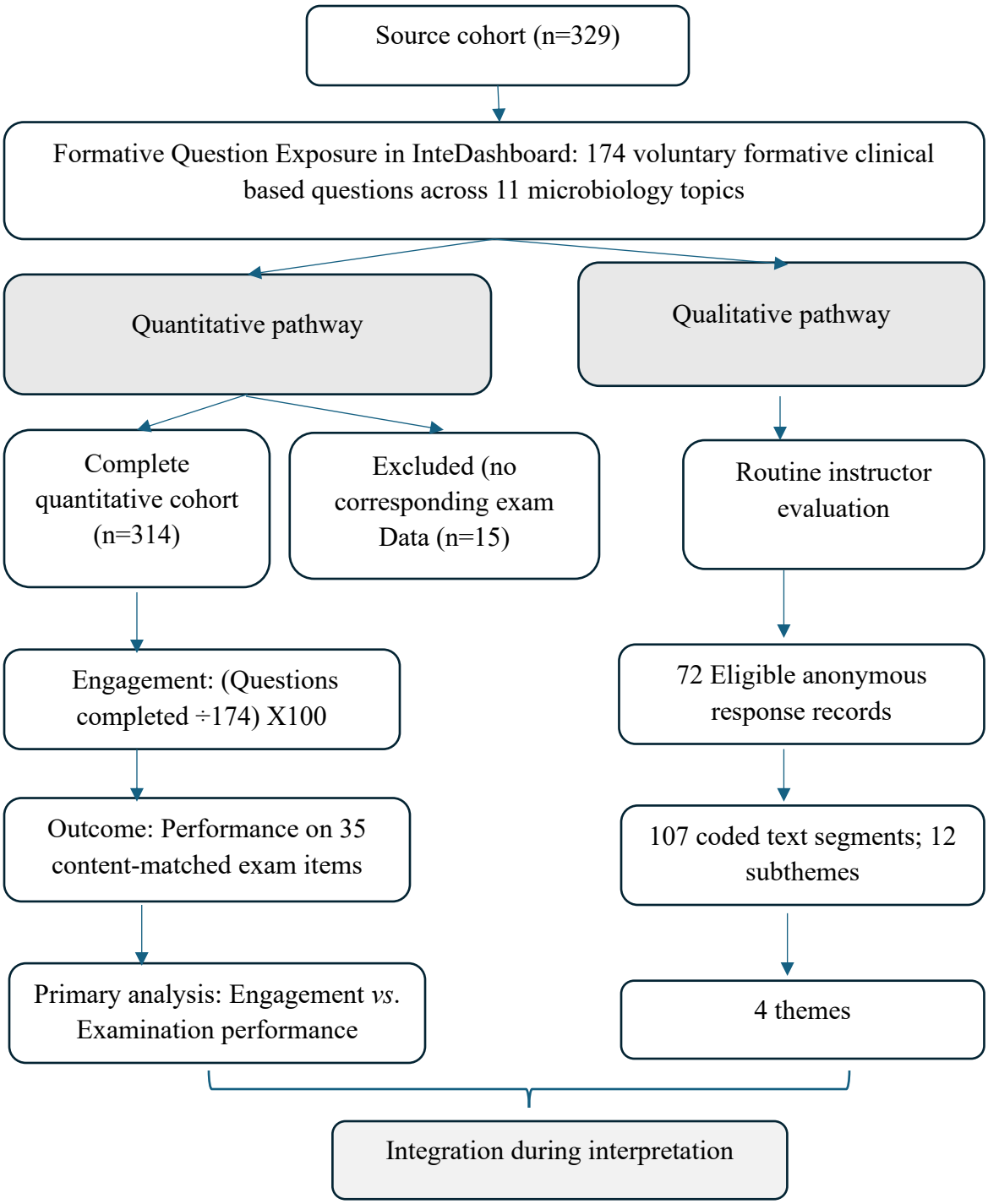


Figure 1: Study design and analytical flow. The quantitative and qualitative components were analyzed separately and integrated during interpretation. Qualitative responses were not linked to individual engagement or examination records.

A. Quantitative Component

Participants

The source cohort comprised all 329 first-year osteopathic medical students enrolled during the study period. Students with complete formative-question engagement data and corresponding examination data were included in the quantitative analysis (n = 314). Fifteen students were excluded because corresponding examination data were unavailable.

Formative Practice Questions

The investigator developed 174 nongraded, clinically oriented multiple-choice formative practice questions aligned with session learning objectives across 11 microbiology topics. All the questions were delivered through InteDashboard (InteDashboard, Inc., USA), multiple attempts were allowed, and explanatory feedback was provided after each attempt. Participation was voluntary and did not contribute to course grades.

Summative examination items

The investigator also developed summative examination items corresponding to the same 11 microbiology topics. The Assessment and Evaluation Center at Burrell College of Health Sciences determined the number of items included in the course assessments; the items were designed to have comparable levels of difficulty. Summative examinations were administered through ExamSoft (ExamSoft Worldwide, Inc., USA) as part of the integrated course assessment. For this study, 35 items representing the 11 topics were identified across the course examinations and aggregated for analysis; they did not constitute a separate assessment. The distribution of formative questions and corresponding examination items is provided in Table 1.

Table 1. Distribution of formative questions and corresponding summative examination items across microbiology topics

Microbiology topics	Formative questions, n	Examination items, n
Normal Flora	20	3
Introduction to Parasites	15	3
Introduction to Fungi	15	2
Helminthic Infections of Solid Tissues	16	3
Filoviruses and Arenaviruses	15	3
Obligate Intracellular Bacteria	12	3
Common Cold	25	6

Gram Positive Rods (*Corynebacterium* and *Listeria*)	15	3
Atypical Bacteria	15	3
Systemic and Dimorphic Fungi	10	3
Mycobacteria and Related Acid Fast Bacteria	16	3
Total	**174**	**35**

***Note**: Each of these topics was covered in a 1-hour lecture, except Common Cold, which was covered in a 2-hour lecture.*

Measures

Formative-Question Engagement

Engagement data were obtained from InteDashboard learning analytics and defined as the number of formative questions completed. The percentage of engagement was calculated as (number of questions completed ÷ 174) × 100. Engagement was analyzed primarily as a continuous variable.

For secondary descriptive analyses, students were grouped according to the empirical distribution of engagement in the analytic cohort: low (≤25th percentile; ≤57.6% completed), moderate (>25th percentile through the median; >57.6% to ≤82.2% completed), and high (>median; >82.2% completed). These categories represented relative engagement within this cohort and were not treated as validated educational thresholds or recommended completion targets.

Examination Performance

The primary outcome was the percentage of the 35 summative examination items answered correctly. Topic-specific performance was calculated as the percentage of corresponding examination items answered correctly within each microbiology topic. Because each topic was represented by only 2–6 examination items, topic-specific analyses were considered exploratory.

Quantitative analysis

Quantitative analyses were performed via IBM SPSS Statistics for Windows, version 31.0 (IBM Corp., Armonk, NY, USA). Continuous variables were summarized using means and standard deviations, and categorical variables were summarized using frequencies and percentages. All the statistical tests were two-sided, with $p < 0.05$ considered statistically significant.

The primary association between formative-question engagement and overall examination performance was examined via Pearson correlation and simple linear regression. The number of formative questions completed was the predictor, and the number of corresponding examination items answered correctly was the outcome. The regression results included the unstandardized

coefficient (B), standard error (SE), 95% confidence interval (CI), standardized coefficient (β), R^2, adjusted R^2, and F statistic. Regression assumptions were assessed via residual plots and the Durbin–Watson statistic.

For secondary analyses, examination performance was compared across the three engagement categories via one-way analysis of variance (ANOVA), followed by Tukey's honestly significant difference (HSD) test for pairwise comparisons. Eta squared (η^2) was used to quantify the overall effect size, whereas Hedges' *g* was calculated to quantify the standardized magnitude of pairwise differences. Topic-specific Pearson correlations were used to examine associations between engagement and examination performance within each of the 11 microbiology topics. P values for these topic-level analyses were adjusted via the Benjamini–Hochberg false discovery rate (FDR) procedure.

B. Qualitative Component

Qualitative Data Source and Eligibility

The qualitative component used deidentified responses to two open-ended questions from the routine instructor evaluation:

1. *"In what ways did the instructor facilitate your learning?"* (Q9 of the evaluation document)
2. *"What suggestions do you have to help your instructor improve their facilitation of student learning?"* (Q10 of evaluation document)

All available responses were reviewed. Responses were eligible if they explicitly referred to the formative practice questions or clearly identifiable question-based resources associated with the study intervention, such as InteDashboard-based practice questions. Comments addressing only general instructor characteristics, lecture delivery, presentation materials, or nonspecific learning resources were excluded from the qualitative analysis.

Qualitative Analysis

Eligible responses were analyzed via inductive thematic analysis. Anonymous response records were the unit of analysis, and meaningful text segments referring to the formative practice questions were the unit of coding. Because a response could contain multiple concepts, it could contribute to more than one code or theme.

Microsoft Copilot Researcher (Microsoft Corporation, Redmond, WA, USA) was used as an AI-assisted tool to support coding and pattern identification in the deidentified responses. The investigator reviewed AI-assisted outputs for accuracy and consistency with the source responses and retained responsibility for coding decisions, theme development, interpretation, and reporting. Independent duplicate human coding, formal consensus coding, and member checking

were not performed. The responses were anonymous and could not be linked across the two evaluation questions or to individual quantitative records.

C. Mixed Methods Integration

Quantitative and qualitative findings were integrated during interpretation at the cohort and educational-resource levels to provide complementary perspectives on formative-question engagement and examination performance. Because qualitative responses were anonymous and could not be linked to individual engagement or examination records, integration was not conducted at the individual-student level.

Ethics approval and consent to participate

The study was reviewed and determined to be exempt by the Burrell College of Osteopathic Medicine Institutional Review Board (BURRELL IRB 0171_2025) on May 12, 2025. The study involved analysis of data generated through routine educational activities, including voluntary engagement with non-graded formative practice questions and course assessment data. The IRB-approved protocol specified that informed consent was not needed for this research; therefore, informed consent to participate was not obtained. Student identifiers were used only for the initial linkage of engagement and examination data and were subsequently removed, with the final analytical dataset de-identified. An amendment adding the qualitative component, based on anonymous qualitative data from course and faculty evaluations, was approved by the same IRB on July 27, 2026.

Results

Participant characteristics

Of the 329 first-year students in the Fall 2025 cohort, 15 were excluded because corresponding examination data were unavailable. The quantitative analysis therefore included 314 students. The students completed a mean of 128.3 ± 47.6 of 174 available formative questions (73.7 ± 27.4%). The mean performance on the 35 content-matched summative examination items was 79.4 ± 11.1%.

Association between Formative Question Engagement and Examination Performance

Pearson correlation analysis demonstrated a positive association between formative question engagement and examination performance ($r = 0.448$, $p < 0.001$).

In simple linear regression, the number of questions completed was significantly associated with the number of examination items answered correctly, $F(1, 312) = 78.36$, $p < 0.001$. The model explained 20.1% of the observed variance in examination performance ($R^2 = 0.201$; adjusted $R^2 = 0.198$). Each additional formative question completed was associated with an estimated 0.0366 additional correct examination items (B = 0.0366, SE = 0.0041, 95% CI 0.0285–0.0448; β = 0.448; $t = 8.85$, $p < 0.001$). According to the fitted model, approximately 27 additional questions corresponded to one additional correct content-matched examination item.

Examination Performance by Engagement Category

Using cohort-derived engagement categories, 79 students were classified as having low engagement (≤57.6% completed), 81 as having moderate engagement (>57.6% to 82.2%), and 154 as having high engagement (>82.2%). The mean examination performance increased across the categories: 71.8 ± 13.0% in the low-engagement group, 78.0 ± 8.9% in the moderate-engagement group, and 84.0 ± 8.5% in the high-engagement group. The between-group difference was statistically significant, $F(2, 311) = 40.53$, $p < 0.001$, with $\eta^2 = 0.207$ (Table 2).

Table 2: Academic Performance by Engagement Category

Engagement Category	Practice question completion range (%)	Number of students (n)	Engagement %, Mean (SD)	Exam Score %, Mean (SD)	Statistical comparison
Low Engagement	≤57.6	79	33.0 ± 16.3	71.8 ± 13.0	ANOVA: $F(2,311) = 40.53$, $p < 0.001$; $\eta^2 = 0.207$
Moderate Engagement	>57.6-≤82.2	81	71.7 ± 6.8	78.0 ± 8.9	
High Engagement	>82.2	154	95.6 ± 5.7	84 ± 8.5	

Note:
Post hoc Tukey comparisons: all pairwise $p < 0.001$;
Hedges' g: low vs moderate = 0.56; low vs high = 1.18; moderate vs high = 0.69

Post hoc Tukey comparisons demonstrated the following:

- High-engagement students scored significantly higher than low-engagement students did (mean difference = 12.2% Exam points, 95% CI: 9.0–15.5, $p < 0.001$; Hedge's g= 1.18).
- High-engagement students scored significantly higher than moderate-engagement students did (mean difference = 6.0% Exam points, 95% CI: 2.8–9.2, $p < 0.001$; Hedge's g= 0.69).
- The moderate-engagement students scored significantly higher than the low-engagement students did (mean difference = 6.2% Exam points, 95% CI: 2.5–9.9, $p < 0.001$; Hedge's g=0.56).

Topic-Specific Analyses

Positive associations between formative-question engagement and topic-specific examination performance were observed across all 11 microbiology topics (Table 3).

Table 3: Associations Between Topic Specific Formative Question Engagement and Academic Performance

Topics	Engagement (%), Mean (SD)	Academic Performance (%), Mean (SD)	Pearson r	Adjusted *p*-values
Normal Flora	83.3 (36.3)	88.2 (19.9)	0.293	<0.001
Introduction to Parasites	79.5 (39.3)	86.0 (20.7)	0.270	<0.001
Introduction to Fungi	84.1 (36.2)	89.8 (21.7)	0.168	0.003
Helminthic Infections of Solid Tissues	58.5 (48.0)	74.0 (28.9)	0.358	<0.001
Filoviruses and Arenaviruses	71.1 (44.6)	77.9 (25.8)	0.230	<0.001
Obligate Intracellular Bacteria	68.0 (45.4)	78.3 (28.3)	0.341	<0.001
Common Cold	75.8 (39.7)	73.9 (18.2)	0.288	<0.001
Gram Positive Rods (*Corynebacterium* and *Listeria*)	69.3 (45.4)	70.3 (26.0)	0.188	<0.001
Atypical Bacteria	77.9 (40.3)	85.2 (25.7)	0.416	<0.001
Systemic and Dimorphic Fungi	73.5 (43.4)	79.7 (27.1)	0.417	<0.001
Mycobacteria and Related Acid Fast Bacteria	65.6 (46.5)	78.3 (26.7)	0.425	<0.001

Note: *N=314; Adjusted p values were calculated using the Benjamini–Hochberg false discovery rate procedure. All reported associations remained statistically significant after correction for multiple comparisons.*

The correlation coefficients ranged from r = 0.168 for Introduction to Fungi to r = 0.425 for Mycobacteria and Related Acid-Fast Bacteria. After Benjamini–Hochberg false-discovery-rate adjustment, all 11 topic-level associations remained statistically significant; the adjusted p values were 0.003 for Introduction to Fungi and <0.001 for the remaining 10 topics.

Qualitative findings

The qualitative analysis included 72 eligible anonymous response records: 54 from the instructor-evaluation question asking how the instructor facilitated learning (Q9) and 18 from the question soliciting suggestions for improving learning facilitation (Q10). Because the responses were

anonymous and could not be linked across prompts, the 72 records did not necessarily represent 72 unique students. The analysis yielded 107 non-mutually exclusive coded text segments organized into 12 refined subthemes and four overarching themes (Table 4). Segment counts documented the analytic process and were not interpreted as participant prevalence.

Table 4. Themes and subthemes identified from student comments about the formative practice questions

Theme	Subthemes	Central finding	Representative excerpt
Assessment readiness and curricular alignment	Examination preparation; assessment representativeness; curricular alignment	Students valued questions that supported examination preparation and reflected taught content, learning objectives, and assessment expectations.	"Very representative of exam questions."
Reinforcement and application of knowledge	General perceived usefulness; reinforcement and understanding; application and cognitive engagement	Students perceived the questions as useful for reinforcing understanding, clarifying difficult concepts, and actively applying microbiology knowledge.	"Help solidify understanding of difficult content material."
Metacognitive calibration and explanatory feedback	Self assessment and calibration; explanatory feedback	Questions and explanations helped students gauge their understanding, identify areas for further study, and guide subsequent learning.	"A representative way to gauge my understanding."
Conditions affecting educational value	Quantity and availability; delivery and accessibility; question-resource quality; difficulty and appropriateness	Students identified access, resource quality, and the perceived difficulty and appropriateness of questions as factors influencing their educational value.	"More practice questions on LEO and Intedashboard would be great, but other than that no comments."

Note: *Responses could contribute to multiple subthemes and themes. Therefore, themes were not mutually exclusive.*

The four themes were (1) assessment readiness and curricular alignment, (2) reinforcement and application of knowledge, (3) metacognitive calibration and explanatory feedback, and (4) conditions affecting educational value.

The students described the questions as useful for preparing for assessments and connecting practice with course content and learning objectives. They also perceived the questions as reinforcing understanding, supporting the application of microbiology concepts, and helping them identify areas requiring further study. Explanatory feedback was described as useful for checking understanding and guiding subsequent learning. These perceptions reflect reported educational value and do not establish an individual-level mechanism linking question or feedback use with examination performance.

Students additionally identified resource conditions that shaped perceived usefulness, including the number and availability of questions, accessibility and delivery, question quality, and perceived difficulty and appropriateness. The comments also included requests for additional practice opportunities. These findings reflect students' reported perceptions and do not establish that any particular feature caused improved examination performance.

Integration of quantitative and qualitative findings

The quantitative and qualitative findings were complementary at the cohort and educational-resource levels. Greater engagement was associated with greater performance on content-matched examination items, whereas students' comments described perceived value in assessment preparation, curricular alignment, knowledge reinforcement and application, self-monitoring, and explanatory feedback. Students also identified resource characteristics that could enhance or constrain usefulness.

Because anonymous qualitative responses could not be linked to individual engagement or examination records in this study, the qualitative findings provide context for, but do not explain, the observed quantitative associations. Integration was therefore conducted at the cohort and educational-resource levels.

Discussion:

In this quantitative-priority observational mixed-methods study, greater voluntary engagement with faculty-developed, clinically oriented formative microbiology questions was moderately associated with higher performance ($r = 0.448$) on content-matched summative examination items. Engagement statistically accounted for 20.1% of the observed variance ($R^2 = 0.201$) in performance. Examination performance also increased progressively across the cohort-derived engagement categories, and positive associations were observed across all 11 microbiology topics. Students perceived the resource as supporting assessment readiness and curricular alignment, reinforcement and application of knowledge, metacognitive calibration and explanatory feedback while identifying conditions that influenced its educational value. Together, these findings identify voluntary formative-question engagement as a meaningful behavioral correlate of content-

matched examination performance. Because engagement was self-selected and the study was observational, the results do not establish that completing more questions resulted in higher examination scores.

Formative-question engagement and examination performance

The positive association between formative-question engagement and examination performance is consistent with the literature on formative assessment, retrieval practice, and self-regulated learning [1–4]. Additional evidence from health-profession education supports test-enhanced learning and repeated retrieval as strategies that can strengthen retention [15, 16]. The present study extends this literature by examining the voluntary use of a formative resource embedded within an authentic preclinical curriculum rather than retrieval practices under controlled conditions. Retrieval-practice evidence therefore provides a plausible educational context for the findings but does not establish the exact mechanism underlying the observed association.

The voluntary nature of engagement is important when interpreting the results. Question completion may reflect not only exposure to formative practice but also differences in prior academic preparation, motivation, self-regulation, attendance, study time, confidence, or use of other learning resources. Thus, R^2=0.201 indicates that engagement statistically accounted for 20.1% of the observed variance in examination performance within this cohort; however, it does not indicate that question completion itself caused 20.1% of performance. Most of the variation remained unexplained by engagement alone.

The graded pattern across engagement categories provides complementary evidence, with progressively higher performance from low to moderate to high engagement. However, these categories were derived from the empirical engagement distribution of this cohort and should not be interpreted as validated thresholds, minimum completion targets, or evidence of a prescribed question "dose." The continuous analysis therefore remains the primary representation of the engagement–performance relationship.

Previous studies have linked engagement with peer-generated and cocreated question banks to summative performance [12–14]. Associations have also been reported for formative assessments in anatomy and medical microbiology [17, 18] and for voluntary engagement with a commercial question bank in relation to standardized examination performance [19]. These contexts differ meaningfully from the present intervention. Peer-generated platforms may combine answering with question authoring, rating, and discussion [13, 14], whereas commercial question banks are not necessarily matched to a specific local curriculum [19]. The present study therefore contributes a distinct perspective by examining faculty-developed, clinically oriented questions aligned with course learning objectives, accompanied by explanatory feedback, and linked to content-matched summative items within the same microbiology curriculum. Such alignment may provide educational advantages by reinforcing intended learning outcomes and strengthening the

coherence between teaching, learning activities, and assessment. Furthermore, institutionally developed resources may offer an accessible formative learning tool without requiring students to purchase additional commercial study materials.

Similarly, prior qualitative studies have shown that medical students use practice questions to prepare for assessments, identify knowledge gaps, guide subsequent studies, and learn from explanatory rationales [20, 21]. Moreover, these studies suggest that how students engage with question banks may shape their educational value, with potentially superficial strategies such as memorization and cueing described alongside concerns about curricular alignment and resource quality [20, 21]. Educational value may therefore depend not only on the amount of engagement but also on how the resource is designed and integrated into the curriculum.

Clinically oriented questions and microbiology learning

Preclinical microbiology requires learners to integrate microbial characteristics, pathogenesis, host responses, clinical manifestations, diagnostic findings and antimicrobial principles. The microbiology and infectious disease education literature supports connecting foundational microbiological knowledge with clinically relevant applications [10, 11]. Clinically oriented formative questions may provide opportunities to retrieve foundational information while considering its relevance within patient-centered problems. This approach is consistent with efforts to introduce application and clinical reasoning earlier in undergraduate medical education [7, 8]. However, clinical reasoning was not directly assessed in the present study. Accordingly, the findings should be interpreted as relating to performance on content-aligned course assessment items rather than as evidence of improved clinical reasoning.

Within an osteopathic medical curriculum, the clinical framing of foundational-science questions is also conceptually compatible with the application-oriented structure of the COMLEX-USA Blueprint [9]. This connection is contextual only. The formative questions in the present study were faculty-developed course items, were not validated COMLEX-USA items, and the study did not evaluate COMLEX-USA readiness or performance. The findings, therefore, should not be interpreted as evidence that engagement in those formative questions certainly improves or predicts licensing-examination performance.

Positive topic-level associations across all 11 microbiology topics after Benjamini–Hochberg false-discovery-rate correction suggest that the overall relationship was not driven by a single content area. However, individual topics contained only two to six summative items. Topic-specific estimates are therefore less precise than the aggregated 35-item outcome and may be influenced by numerous factors, such as item characteristics, content complexity, and differences in student preparation. These analyses should be considered supportive and exploratory rather than evidence of equal benefit across topics or greater effectiveness in particular topics.

Potential educational mechanisms

Several processes could plausibly contribute to the observed association, although none was directly tested in the present study as a mediator. Answering formative questions requires the retrieval of previously learned information, and repeated retrieval has been associated with improved retention [2, 3, 15, 16]. Clinically contextualized questions may provide opportunities to apply foundational knowledge within patient-oriented problems [7, 8]. Practice questions may also support self-regulated learning by helping learners recognize uncertainty, identify knowledge gaps, and direct subsequent study [4, 22], although learners may not always accurately assess their own knowledge [23]. Finally, explanatory rationales may extend the value of formative questions beyond indicating whether an answer is correct. Feedback that explains errors and relevant concepts can support learning by helping students understand why an answer is correct or incorrect [5, 6, 24].

These mechanisms remain plausible interpretations rather than demonstrated explanations. The study did not measure retrieval strength, number of attempts, response accuracy, time on task, spacing of practice, explanation use, metacognitive calibration, or subsequent study behavior. Long-term retention, transfer to novel problems, and clinical reasoning were also not assessed. Accordingly, the findings demonstrate an association between behavioral engagement with formative resources and content-matched examination performance but do not establish mediation through any specific cognitive mechanism.

Student perceptions and educational value

The four qualitative themes in the current study-assessment readiness and curricular alignment, reinforcement and application of knowledge, metacognitive calibration and explanatory feedback, and factors influencing educational value-are consistent with previously reported perspectives on medical students' use of practice questions [20, 21].

Students particularly appreciated practice questions that were aligned with lecture content, stated learning objectives, and assessment expectations. This perception is consistent with the study's deliberate alignment of formative questions with session learning objectives and the use of content-matched summative examination items as the quantitative outcome. This alignment reflects the principle of constructive alignment, which emphasizes coherence among learning outcomes, instructional activities, and assessments [25]. Nevertheless, students' perceptions of assessment representativeness should not be interpreted as evidence of the psychometric equivalence between formative and summative questions.

Students also reported that practice questions helped reinforce and apply knowledge, reflecting principles of retrieval-based learning [3, 15] and the application of foundational knowledge within clinically contextualized problems [8]. Their descriptions of using questions to identify uncertainty and guide further study were consistent with self-regulated learning processes [4, 22]. The importance students place on explanatory feedback is likewise consistent with evidence that

providing explanations can support learning and transfer beyond simply indicating whether an answer is correct [5, 6, 24]. However, these qualitative perceptions provide plausible accounts of how formative-question engagement may support learning but do not establish the cognitive mechanisms underlying the observed quantitative association with examination performance. Moreover, although self-assessment has recognized limitations [23], the present study did not capture whether students accessed, revisited, or used the provided explanations to modify their subsequent learning activities. Thus, while explanatory feedback may contribute to the educational value of formative questions, its specific contribution to the association between formative-question engagement and examination performance remains uncertain.

Students further identified quantity, availability, accessibility, coverage, quality, difficulty and appropriateness as conditions affecting educational value. Similar considerations have been reported in qualitative studies examining medical students' use of question banks [20, 21]. Collectively, these findings suggest that educational value may depend not only on the quantity of available questions but also on resource design, alignment with curricular objectives, accessibility, and the quality of accompanying feedback.

Integration of quantitative and qualitative findings

The quantitative and qualitative components provide complementary but distinct perspectives. Greater voluntary engagement was consistently associated with higher content-matched examination performance, whereas students described perceived benefits related to assessment preparation, reinforcement and application, metacognitive monitoring, and explanatory feedback.

Because anonymous qualitative records could not be linked to individual engagement or examination data, the qualitative findings cannot explain why particular students engage more extensively or perform better, nor can they establish mechanisms underlying the quantitative association. Rather, the qualitative findings provide a complementary context at the cohort and resource levels, reflecting the level of integration permitted by the available data [26].

The principal contribution of this study, therefore, is not to establish causal effectiveness but to characterize the association between voluntary engagement with faculty-developed, clinically oriented formative microbiology questions and performance on explicitly content-matched summative examination items while adding students' perceptions of the same resource.

Strengths and limitations

This study has several strengths. Engagement was measured via objective learning platform data rather than self-reports, and formative questions were explicitly matched to summative items by microbiology content. Engagement was analyzed continuously in the primary analysis, with cohort-derived categories and standardized effect sizes providing complementary descriptive information. Topic-level analyses examined consistency across 11 microbiology topics, with Benjamini–Hochberg false-discovery-rate correction. The mixed methods design also incorporates

learner perspectives while maintaining the distinction between unlinked quantitative and qualitative datasets.

The principal limitation of the present study is the observational design and voluntary nature of question use, which precludes causal inference. Students with different engagement levels may have differed in prior achievement, motivation, self-regulation, attendance, study time, confidence, or use of external resources. Because these factors were not adequately measured or controlled, self-selection and residual confounding remain plausible explanations for the observed associations.

The study involved a single cohort at one osteopathic medical school within a specific preclinical microbiology curriculum, limiting generalizability to other institutions, disciplines, and learner populations. Although curricular alignment strengthened the relevance of the locally developed questions, the outcome comprised only 35 content-matched summative items and did not assess comprehensive microbiology competence, long-term retention, transfer, or clinical reasoning. Topic-level analyses were further constrained by two to six summative items per domain, limiting precision and increasing susceptibility to item-specific characteristics.

In this study, engagement was defined as the percentage of available questions completed. This objective measure did not capture attempts, response accuracy, time on task, practice spacing, explanation use, guessing, collaboration, or subsequent study behaviors. More comprehensive learning analytics could provide a richer characterization of engagement. Cohort-derived engagement categories should also be interpreted as descriptive rather than validated educational thresholds.

The qualitative component relied on anonymous written responses from routine instructor evaluations, limiting opportunities for probing and clarification compared with interviews or focus groups [27]. The 72 eligible response records did not necessarily represent 72 unique students, and the 107 non-mutually exclusive coded segments represent units of qualitative content rather than participant prevalence.

AI-assisted tools supported coding organization and pattern identification, whereas the investigator retained responsibility for coding decisions, theme development, interpretation, and reporting. Independent duplicate human coding, formal consensus coding, and member checking were not performed. Finally, anonymity prevents linkage across evaluation questions and between qualitative responses, formative-question engagement, and examination performance, restricting mixed-method integration to the cohort and educational-resource levels [26].

Future directions

Future studies should examine whether these associations are reproducible across institutions, cohorts, foundational-science disciplines, and curricular settings. Whereas feasible, randomized or stronger quasi-experimental designs should measure and account for prior achievement, motivation, self-regulation, attendance, study habits, and the use of other learning resources. Such

designs would help distinguish the contribution of formative-question engagement from learner characteristics associated with both engagement and examination performance.

Learning analytics should extend beyond completion rates to capture attempts, response accuracy, practice timing and spacing, time on task, explanation use, and subsequent study behaviors. Mechanism-focused studies could investigate how clinical contextualization, repeated retrieval, feedback type, and feedback timing influence learning, using outcomes that include delayed retention and transfer to novel clinical problems.

Research-specific interviews or focus groups could further examine why students adopt different patterns of question use, how they interact with explanatory feedback, and how practice questions influence subsequent study decisions, extending prior qualitative work [20, 21]. Rigorously designed qualitative questionnaires may offer an alternative when interviews are impractical [27]. Where feasible, linking qualitative perspectives with objective engagement and performance data could enable more comprehensive individual-level mixed-method integration [26].

Conclusion

Among first-year osteopathic medical students, greater voluntary engagement with faculty-developed, clinically oriented formative microbiology questions was consistently associated with higher performance on content-matched summative examination items. Students perceived the resource as supporting assessment readiness and curricular alignment, reinforcement and application of knowledge, metacognitive calibration, and explanatory feedback while identifying conditions that influenced its educational value.

Because engagement was self-selected and potentially important learner characteristics were unmeasured, these findings demonstrate associations rather than effectiveness. They do not establish improvements in clinical reasoning, long-term retention, transfer of learning, or COMLEX-USA performance. Nevertheless, the consistency of the quantitative associations and complementary learner perspectives highlights the potential educational value of thoughtfully designed, clinically oriented, curriculum-aligned formative-question resources in preclinical medical education.

Abbreviations

AI: artificial intelligence
ANOVA: analysis of variance
CI: confidence interval
COMLEX-USA: Comprehensive Osteopathic Medical Licensing Examination of the United States
FDR: false discovery rate
HSD: honestly significant difference
IRB: Institutional Review Board
SD: standard deviation

SE: standard error
SPSS: Statistical Package for the Social Sciences

Declarations

Ethics approval and consent to participate

The study was reviewed by the Burrell College of Osteopathic Medicine Institutional Review Board and determined to be exempt from further review (IRB #0171_2025). All procedures were conducted in accordance with applicable institutional policies and ethical guidelines.

Consent for publication: Not applicable

Availability of data and materials

Deidentified data are available from the author upon reasonable request.

Competing interests

The author declares that there are no competing interests.

Funding

This study received no specific grant or fund from any funding agency.

Authors' contributions

HPN conceptualized the study, developed the methodology, collected and analyzed the data, interpreted the findings, and drafted and revised the manuscript. HPN read and approved the final manuscript.

Acknowledgments

The author extends sincere gratitude to the students whose participation and feedback contributed to this study. The author also gratefully acknowledges the Compliance Office and Academic Affairs at Burrell College of Health Sciences for the institutional support and the College's Evaluation and Assessment Office for providing examination-related data.

Declaration on the use of generative AI tools

Microsoft Copilot Researcher was used as an assistive tool for preliminary coding and pattern identification of deidentified qualitative responses. All coding outputs, thematic structures, and interpretations were independently reviewed, verified, and finalized by the investigator. OpenAI ChatGPT was also used for language editing, clarity, and formatting suggestions. No generative AI system independently interpreted the findings, made scientific judgments, or determined the conclusions of the study. The author retains full responsibility for the accuracy, integrity, and final content of the manuscript.

References

1. Lipnevich AA, Mattern K, Feddock C. Formative assessment and feedback in medical education: a practical guide: AMEE Guide No. 189. Med Teach. 2026;48:921–940. doi:10.1080/0142159X.2025.2569623.
2. Roediger HL, Karpicke JD. Test-enhanced learning: taking memory tests improves long-term retention. Psychol Sci. 2006;17:249–255. doi:10.1111/j.1467-9280.2006.01693.x.
3. Trumble E, Lodge J, Mandrusiak A, Forbes R. Systematic review of distributed practice and retrieval practice in health professions education. Adv Health Sci Educ Theory Pract. 2024;29:689–714. doi:10.1007/s10459-023-10274-3.
4. Zheng B, Sun T. Self-regulated learning and learning outcomes in undergraduate and graduate medical education: a meta-analysis. Eval Health Prof. 2025;48:430–450. doi:10.1177/01632787241288849.
5. Butler AC, Godbole N, Marsh EJ. Explanation feedback is better than correct answer feedback for promoting transfer of learning. J Educ Psychol. 2013;105:290–298. doi:10.1037/a0031026.
6. Levant B, Zückert W, Paolo AM. Postexam feedback with question rationales improves retest performance of medical students on a multiple-choice exam. Adv Health Sci Educ Theory Pract. 2018;23:995–1003. doi:10.1007/s10459-018-9844-z.
7. Specian Junior FC, Santos TM, Sandars J, Amaral EM, Cecilio-Fernandes D. Identifying the response process validity of clinical vignette-type multiple choice questions: an eye-tracking study. Med Teach. 2023;45:845–851. doi:10.1080/0142159X.2023.2182662.
8. Hawks MK, MacIuba JM, Merkebu J, Durning SJ, Mallory R, Arnold MJ, et al. Clinical reasoning curricula in preclinical undergraduate medical education: a scoping review. Acad Med. 2023;98:958–965. doi:10.1097/ACM.0000000000005197.
9. National Board of Osteopathic Medical Examiners. COMLEX-USA blueprint: revised February 2025 [Internet]. Chicago, IL: National Board of Osteopathic Medical Examiners; 2025 [cited 2026 Oct 3]. Available from: https://www.nbome.org/app/uploads/2025/05/COMLEX-USA-Blueprint-2025.pdf
10. Southwick F, Katona P, Kauffman C, Monroe S, Pirofski LA, Del Rio C, et al. Commentary: IDSA guidelines for improving the teaching of preclinical medical microbiology and infectious diseases. Acad Med. 2010;85:19–22. doi:10.1097/ACM.0b013e3181c485c5.
11. Melber DJ, Teherani A, Schwartz BS. A comprehensive survey of preclinical microbiology curricula among US medical schools. Clin Infect Dis. 2016;63:164–168. doi:10.1093/cid/ciw262.

12. Hwang R, Paley M, Bull D, Omrani O, Vergara-Jalandoni D, Egbury G. A systematic review and meta-analysis on the impact of peer-made MCQ question-bank usage on summative assessments in medical education. J Adv Med Educ Prof. 2025;13:259–269. doi:10.30476/jamp.2025.104780.2084.

13. Walsh JL, Harris BHL, Denny P, Smith P. Formative student-authored question bank: perceptions, question quality and association with summative performance. Postgrad Med J. 2018;94:97–103. doi:10.1136/postgradmedj-2017-135018.

14. Guilding C, Pye RE, Butler S, Atkinson M, Field E. Answering questions in a cocreated formative exam question bank improves summative exam performance, while students perceive benefits from answering, authoring, and peer discussion: a mixed methods analysis of PeerWise. Pharmacol Res Perspect. 2021;9:e00833. doi:10.1002/prp2.833.

15. Green ML, Moeller JJ, Spak JM. Test-enhanced learning in health professions education: a systematic review: BEME Guide No. 48. Med Teach. 2018;40:337–350. doi:10.1080/0142159X.2018.1430354.

16. Larsen DP, Butler AC, Roediger HL 3rd. Repeated testing improves long-term retention relative to repeated study: a randomized controlled trial. Med Educ. 2009;43:1174–1181. doi:10.1111/j.1365-2923.2009.03518.x.

17. Kingston AK, Garofalo EM, Cardoza K, Fisher RE. Designing formative assessments to improve anatomy exam performance. Anat Sci Educ. 2023;16:989–1003. doi:10.1002/ase.2279.

18. Imwattana K, Ngamskulrungroj P. Comparison of online quizzes and standard summative examination for the evaluation and guidance of students in medical bacteriology and mycology. BMC Med Educ. 2025;25:342. doi:10.1186/s12909-025-06926-0.

19. Sherman EC, Stomberger JI, Porter-Stransky KA. Association between ScholarRx Qmax question bank engagement and NBME CBSE and USMLE Step 1 performance. BMC Med Educ. 2026;26:1453. doi:10.1186/s12909-026-10013-3.

20. Sheehy R, Scott D, Davis D, Roffler M, Sweatman T, Nemec E. Medical student use of practice questions in their studies: a qualitative study. BMC Med Educ. 2024;24:1181. doi:10.1186/s12909-024-06168-6.

21. Fisher J, Leahy D, Lim JJ, Astles E, Salvatore J, Thomson R. Question banks: credit? Or debit? A qualitative exploration of their use among medical students. BMC Med Educ. 2024;24:569. doi:10.1186/s12909-024-05517-9.

22. Andrews MA, Okuliar CA, Whelton SA, Windels AO, Kruse SR, Nachnani MG, et al. Using self-regulated learning microanalysis to examine regulatory processes in clerkship

students engaged in practice questions. Perspect Med Educ. 2023;12:385–398. doi:10.5334/pme.833.

23. Yates N, Gough S, Brazil V. Self-assessment: with all its limitations, why are we still measuring and teaching it? Lessons from a scoping review. Med Teach. 2022;44:1296–1302. doi:10.1080/0142159X.2022.2093704.

24. Ryan A, Judd T, Swanson D, Larsen DP, Elliott S, Tzanetos K, et al. Beyond right or wrong: more effective feedback for formative multiple-choice tests. Perspect Med Educ. 2020;9:307–313. doi:10.1007/s40037-020-00606-z.

25. Biggs J. Enhancing teaching through constructive alignment. High Educ. 1996;32:347–364. doi:10.1007/BF00138871.

26. Battista A, Torre D, Konopasky A. Essential concepts for effective mixed methods research in the health professions: AMEE Guide No. 173. Med Teach. 2025;47:766–778. doi:10.1080/0142159X.2024.2401464.

27. Tombs M, Strange H. Using qualitative questionnaires in medical education research. Perspect Med Educ. 2024;13:280–287. doi:10.5334/pme.1102.