European Journal of Dental Education

WILEY

ORIGINAL ARTICLE OPEN ACCESS

The Role of Chatbots in Enquiry-Based Learning for Oral Health Students—An Exploratory Study

Rohini Khareedi | Daniel Fernandez

Department of Oral Health, School of Clinical Sciences, Auckland University of Technology, Auckland, New Zealand

Correspondence: Rohini Khareedi (rohini.khareedi@aut.ac.nz)

Received: 17 July 2024 | **Revised:** 5 March 2025 | **Accepted:** 27 April 2025

Funding: The authors received no specific funding for this work.

Keywords: Chatbots | enquiry | learning

ABSTRACT

Objective: This study explores the reliability of four Chatbots in enquiry-based learning. Four Chatbots, namely Microsoft Copilot, Google Gemini, ChatGPT 3.5 and Perplexity, were used to answer and generate questions in four specific subject areas.

Methods: The four Chatbots were subjected to questions at three cognitive levels and to generate questions based on specific contexts. The responses generated were assessed by two oral health academics for accuracy and appropriateness.

Results: The findings indicated that ChatGPT3.5 generated the best self-assessment questions while Microsoft Copilot generated the best answers to questions. The performance of the Chatbots varied based on the subject on which the question was based and on the cognitive level of the question. While the questions at cognitive level one were answered most appropriately, the overall depth of responses to periodontology questions at cognitive level two was lower than for questions on dental materials, restorative dentistry and oral biology.

Conclusion: The potential role of Chatbots in enquiry-based learning is evident to some extent, but they currently do not have the proficiency that a human teacher has.

1 | Introduction

Enquiry-based learning is an approach that requires learners to probe actively into problems to understand situations and solve problems. Traditionally, the role of the teacher has been to support the enquiry by stimulating the learner's curiosity, encouraging the learner to question and empowering the learner to evaluate the outcomes before reaching conclusions [1]. The role of the teacher is challenging as learners have different levels of knowledge and each learner processes information at different rates. This can be especially challenging in large classrooms where personalised attention to learners is impossible to achieve. In reality, learners cannot undertake enquiry without guidance to sense the context, understand the problem and hypothesise tentatively [2].

The emergence of generative Artificial Intelligence Chatbots that can replicate human conversation and virtually respond to anything asked in a few seconds at anytime and anywhere has opened their potential as facilitators in enquiry-based learning. Chatbots are being used to answer questions, deliver content, assess learner performance and provide administrative services and support like making announcements on schedules or following up on yet-to-complete compulsory tasks [3].

Though Chatbots' history dates to several decades ago, they have evolved and continue to evolve. Since the introduction of ELIZA, one of the earliest Chatbots, introduced in 1966 by Joseph Weizenbaum, there are more sophisticated and capable Chatbots today with greater access to data. The utility of Chatbots has infiltrated all walks of life, including the

Education sector, with the promise to bridge gaps in teaching and learning by creating personalised content for individual learners and better learner engagement with realistic simulations and creative tasks while providing answers to questions and generating questions from uploaded data based on the context in real time [4]. A recent 2024 study has assessed the potential ability of Chatbots to be integrated into dental education programmes for automated feedback for student assessment of dental radiographs [5].

However, to integrate Chatbots into enquiry-based learning, it is important to assess their credibility and reliability. Published research on the ability of Chatbots has uncovered variations in performance, including inaccuracies based on the questions' subject and the cognitive level of the questions asked [6–8]. Chatbots can sometimes produce misinformation through confabulations and may exhibit bias due to their dependence on skewed data [9].

The purpose of this study was to explore the reliability of four Chatbots to generate questions based on specific contexts and respond to questions at three different cognitive levels in four specific subject areas.

2 | Methods

This study is nonhuman subject research using publicly available Chatbots and hence did not require ethical approval. The four chatbots—Microsoft Copilot, Google Gemini, ChatGPT 3.5 and Perplexity were—selected based on anecdotal evidence suggesting they are the most frequently used by our students. The responses to the prompts from the four Chatbots were assessed by two academics in the Oral Health department. The two academics are authors of this paper and have been teaching faculty in the department of Oral Health for over 15 years. One academic is a registered dentist while the other is a registered dental therapist. Both academics have postgraduate qualifications. All prompts required a text-based response. The responses were scored on a numerical scale of 0–5 by the two academics separately and a decision was made to use a third assessor if the allocated scores varied by more than 1 point. The final score was the average of the scores given by the two academics. The scoring was based on the expectations of second-year undergraduate students enrolled in the Bachelor of Health Science in Oral Health. The rubric used for scoring is depicted in Appendix A.

The four Chatbots were used to generate questions from a hypothetical clinical scenario and an open-access journal publication. The publication was downloaded from the site and uploaded to each of the Chatbots to elicit a response. The prompt used for the hypothetical clinical scenario was: Generate four questions for consideration in this clinical context and the prompt used for the article was: Summarise the article and generate five self-assessment questions. The quality and appropriateness of the four questions to the clinical context and the self-assessment questions generated were scored on a numerical scale of 0–5, with a higher score indicating superior performance. The summary of the article generated by each of the Chatbots was scored for comprehensiveness and clarity. The four chatbots were also evaluated on their ability to facilitate a guided inquiry.

Two oral health academics retrieved questions and scenarios from the peer-reviewed question banks that the faculty have developed for student learning and categorised the questions based on the cognitive level into three levels, with questions at level 1 requiring mostly remembering, while questions at level 2 requiring understanding, analysing and comparing, and level 3 questions tested the ability to apply knowledge to a specific context.

- Level 1 included questions that began with the words: 'name', 'list', 'what' and 'which'. A total of 40 questions from periodontology, oral biology, restorative dentistry and dental materials were chosen.
- Level 2 included questions that began with the words: 'describe', 'discuss', 'correlate' 'explain' and 'compare'. A total of 20 questions from periodontology, oral biology restorative dentistry and dental materials were chosen.
- Level 3 questions were 'application' based on five clinical contexts.

The list of questions used as prompts for all three levels is included in Appendices B–D.

The responses were scored on a numerical scale of 0–5 by the two academics. A single score was allocated for level 1 and 3 questions for accuracy. Responses to level 2 questions were scored for accuracy, depth and clarity separately. The average of the scores allocated by the two academics was the final score for each response. A third assessor was not needed as the scores allocated by the two academics did not differ by more than a point.

A comparative analysis was done using descriptive statistical tests to compare the performance of the four selected Chatbots. The tests were carried out using Microsoft Excel. Inferential statistical tests were not conducted.

3 | Results

The questions generated for a hypothetical clinical problem by Microsoft Copilot, Google Gemini, ChatGPT 3.5 and Perplexity are presented in Table 1. All four Chatbots covered most of the points expected. However, the response from Microsoft Copilot and Gemini did not include an explicit reference to dental caries.

The summary of the open-access journal article on-Caries Management by Risk Assessment published in 2018 in Advances in Dental Research [10] and the self-assessment questions generated were assessed and scored as shown in Figure 1. Chat GPT 3.5 was scored 5 for clarity and comprehension of the summarised text, while Gemini and Perplexity were allocated a score of 3.5 for comprehensiveness. The score received by Gemini for appropriate self-assessment questions generated was the lowest at 3.5. Microsoft Copilot and Perplexity both received a score of 4.5.

All four Chatbots were asked the same question and based on the response the key words were used to generate

TABLE 1 | Questions generated for the following hypothetical scenario: A 5-year-old has a swollen upper lip and pain in the upper left central incisor.

Microsoft Copilot	ChatGPT3.5	Google Gemini	Perplexity
1. What is the cause of the swelling? 2. Is there a history of allergies or recent exposure to allergies? 3. Is there any dental trauma or infection? 4. Are there any other associated symptoms?	1. What could be the potential causes of the swollen upper lip and pain in the upper left central incisor in a 5-year-old child? 2. Has the child experienced any recent trauma or injury to the mouth or face that could explain the symptoms? 3. Are there any signs of dental caries or infection in the affected tooth upon examination? 4. Does the child have a history of dental issues or any underlying medical conditions that may contribute to the symptoms?	1. Has the swelling gotten worse since it started? Is there any trouble breathing or swallowing? These could be signs of a more serious allergic reaction and require immediate medical attention. 2. Has your child recently fallen or bumped their mouth? Have they eaten anything new or unusual that could cause an allergic reaction? Knowing the potential cause can help determine the best course of action. 3. How much pain is your child in? Are they able to eat and sleep comfortably? 4. Is the tooth loose or discoloured? Has there been any bleeding or pus? These could be signs of a dental infection, and a dentist visit may be necessary.	1. Has the child experienced any recent trauma or injury to the mouth or face area that could have caused the swollen lip and tooth pain? 2. Is there any visible sign of tooth decay, discolouration, or abscess formation around the affected upper left central incisor? 3. How long has the swelling and pain been present, and has it been gradually worsening or fluctuating in severity? 4. Does the child have any history of dental issues, such as previous tooth infections or abscesses, which could be related to the current symptoms?

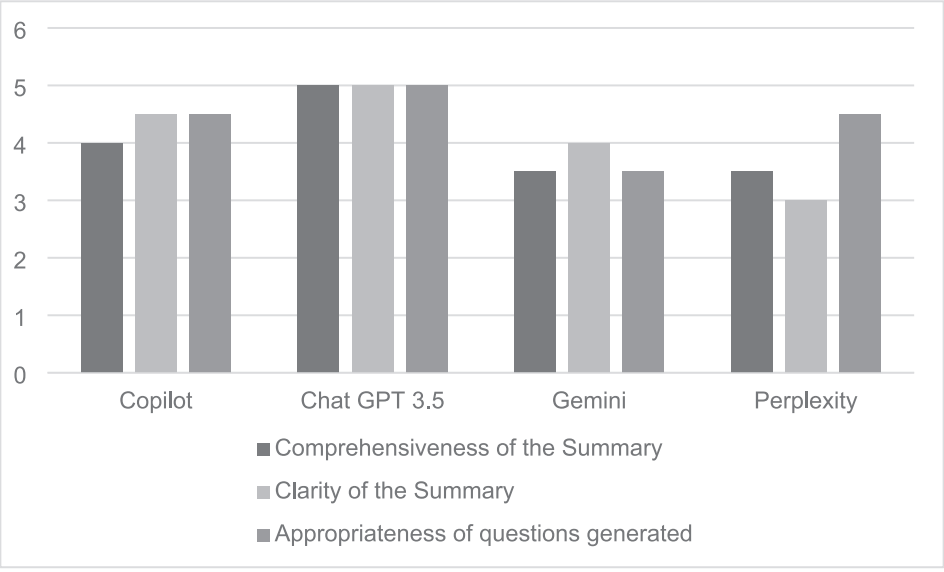

FIGURE 1 | Scores of Chatbots for summarising articles and generating questions.

further questions as highlighted in Table 2. As shown in Table 2, the responses from all Chatbots eventually were able to generate the question on acceptable platelet counts for dental treatment in HIV-positive patients and thus supported guided enquiry.

The aggregate of the scores by each Chatbot in each of the four subject areas for cognitive level 1 questions was recorded as shown in Table 3. The responses by Perplexity for questions on Dental Materials scored 4.05, while the score allocated was 4.9 for questions on periodontology.

TABLE 2 | Guided enquiry with the four Chatbots.

Copilot	ChatGPT 3.5	Gemini	Perplexity
What should be checked in the medical history of an HIV-positive patient prior to dental treatment? Follow-up question to Question 1: What are the relevant laboratory test values that should be assessed? Follow-up question to Question 2: How does platelet count impact dental treatment? Follow-up to Question 3: What is the acceptable platelet count in a HIV-positive patient for dental treatment?	What should be checked in the *medical history* of an HIV-positive patient prior to dental treatment? Follow-up question to Question 1: How does being HIV positive affect the *bleeding and clotting*? Follow-up question to Question 2: What is *thrombocytopenia*? Follow-up to Question 3: What is *the acceptable platelet count* in a HIV-positive patient for dental treatment?	What should be checked in the *medical history* of an HIV-positive patient prior to dental treatment? Follow-up question to Question 1: Why is there a risk for *bleeding* in HIV-positive patients? Follow-up question to Question 2: What is *thrombocytopenia*? Follow-up to Question 3: What is the *acceptable platelet count* in a HIV-positive patient for dental treatment?	What should be checked in the *medical history* of an HIV-positive patient prior to dental treatment? Follow-up question to Question 1: Why is there a risk for *bleeding* in HIV-positive patients? Follow-up question to Question 2: What is the *acceptable platelet count* in a HIV-positive patient for dental treatment?

The responses were scored for accuracy, depth and clarity as shown in Table 4 for the 20 questions at cognitive level 2. The scores for the depth of the responses were further analysed by subject. The scores allocated for the depth of the responses to the periodontology questions by the four Chatbots were lower than for the questions in the other three subject areas as shown in Figure 2. The scores allocated for the depth of the responses by all four Chatbots to questions on dental materials were the highest.

The responses to the five application-based cognitive level 3 questions by Microsoft Copilot scored the highest while the responses from ChatGPT 3.5 scored the lowest as shown in Table 5.

Examples of responses to prompts and average score allocated is in Appendix E.

4 | Discussion

Enquiry-based learning has been accepted for its learner-centred approach that is driven by questions to build on what is known. Though this inductive approach helps in building the ability to think critically and solve problems, its success depends on the student. The learning process must be engaging while allowing for exploration. The knowledge obtained should be stimulatory for further probing and evaluation. In traditional enquiry-based learning, the teacher guides the enquiry process. However, large classrooms with diverse student academic, professional and social backgrounds challenge this approach [11].

The emergence of artificial intelligence tools like Chatbots that are rapidly evolving has features that enable their integration into education systems for enquiry-based learning. The ability of Chatbots to process natural language allows engagement with learners by replicating human-like conversations [12]. Chatbots can not only generate responses to prompts and questions but also evaluate learning by generating questions. Despite the capabilities of the Chatbots, there have been questions raised about their dependability to always deliver appropriate responses [1, 13]. Additionally, research by Seth et al. suggests variations in Chatbot responses [14].

The four Chatbots used in our study to generate questions to consider for a clinical context showed clear evidence of the ability of all four Chatbots to understand the question, grasp the context and provide a response. This ability is like what was recognised as a feature of Chatbots in a case study on the impact of ChatGPT on software engineering education [12].

Chatbots' ability to improve research efficiency by summarising articles has been well documented [15]. When asked to summarise an article published in the on-Caries Management by Risk Assessment published in 2018 in Advances in Dental Research journal [10], the performance of the four Chatbots was varied, with ChatGPT 3.5 scoring the highest for clarity and comprehensiveness. This finding contrasted with a 2023 study on Rhinoplasty which concluded that Google Bard had more comprehensive responses than ChatGPT and Bing AI [14].

TABLE 3 | Scores for responses to cognitive level 1 questions.

	Copilot	ChatGPT 3.5	Gemini	Perplexity
Oral biology	4.85	4.9	4.9	5
Periodontology	4.5	4.825	5	4.9
Restorative dentistry	4.6	4.85	4.75	4.45
Dental materials	4.85	4.5	4.45	4.05

Chatbots can generate questions from text input that are coherent and contextually relevant. When asked to generate questions based on the 2018 publication on-Caries Management by Risk Assessment, all four Chatbots generated questions, with ChatGPT 3.5 generating the most appropriate questions. However, all the questions generated were at the lower end of Bloom's level, similar to the findings in the study on the ability of ChatGPT to automatically generate educational assessment questions [16].

Chatbots are known to have conversational artificial intelligence that can enhance the learning experience through an interactive and dynamic experience, which is the basis of guided enquiry in enquiry-based learning. When asked to respond to the factors that need to be considered in an HIV-positive patient requiring dental treatment, the responses generated by the Chatbots allowed for additional questions, and those responses allowed further questions. This application of Chatbots indicates that it is a potential tool in higher education. A randomised clinical trial on using artificial intelligence to provide feedback to students on radiographic diagnostic competencies has shown the potential to use Chatbots to enhance student performance [5].

The performance of the four Chatbots varied based on the cognitive level of the questions asked in our study. This finding is similar to the findings in a study comparing the responses of ChatGPT and Google Bard to MCQ-based questions at three cognitive levels [8]. Our findings were also similar to the findings where five different Chatbots were asked for responses to the questions in the medical licence examination [17]. However, for questions based on analysis and application, the performance of Copilot and Gemini was better than ChatGPT 3.5 and Perplexity. These differences could be due to differences in where the data are sourced from, and this finding supports the conclusion by the authors of the study on responses to controversial questions in oral pathology, oral medicine and oral radiology [18]. The authors suggested the importance of collaborating with experts in the field when algorithms are designed for Chatbots.

TABLE 4 | Scores for responses to cognitive level 2 questions.

	Copilot	ChatGPT 3.5	Gemini	Perplexity
Accuracy	4.77	4.9	4.82	4.85
Depth	4.2	4.62	4.23	4.12
Clarity	4.82	4.82	4.9	4.85

TABLE 5 | Scores for responses to the five cognitive level 3 questions.

Copilot	Chat GPT 3.5	Gemini	Perplexity
5	2.5	5	5
5	4	5	3.5
5	5	5	5
4.5	4.5	5	5
5	5	4	5

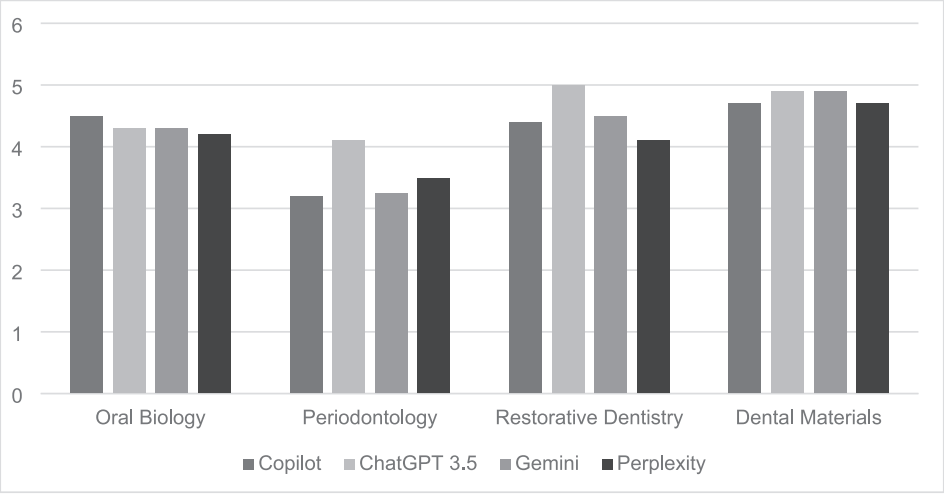

FIGURE 2 | Scores of Chatbots for depth of responses to cognitive level 2 questions.

The Chatbots scored particularly low when assessed for the depth of the answers which were more obvious in responses to questions on periodontology and oral biology. A similar finding was reported in a study on questions on rhinoplasty where the authors acknowledged the limited ability of Chatbots to deliver in-depth responses [14]. There were errors in the responses by all the four Chatbots and this finding is not different from findings in several other studies. The authors of the study using ChatGPT for seeking responses to questions within the scope of the United States Medical Licensing Examination reported inaccuracies [19]. So also the study on Cancer Society's questions to Chat GPT and Bing reported extensive short comings [20].

When the scores of all three cognitive level questions were combined, Copilot had the highest while Perplexity had the lowest. This finding contrasts with the finding of Google Bard performing better than ChatGPT 3.5 and Microsoft Bing in a study on information on rhinoplasty [14] and with the study on lung cancer questions which found ChatGPT 3.5's performance to be better than Microsoft Bing and Google Bard's performances [7].

The performance of the Chatbots varied based on the subject on which the question was based. The overall depth of responses for periodontology questions was lower than for questions on dental materials, restorative dentistry and oral biology. This finding of variation in performance based on the subject is similar to the finding in a study on medical questions where the researchers reported that responses for questions on neurology were better than those on genetics [8].

5 | Limitations

This exploratory study is using Chatbots that are in the public domain and are free to use. Using more advanced Chatbots or using varied prompts may have yielded different results. The selective bias and wording of questions may have affected the responses. Additionally, the scoring by the two oral health academics may have been influenced by implicit factors such as language, interpretation and the style of responses. The analysis of the findings is with descriptive statistics. A mixed-method study with inferential statistics and language analysis is warranted.

6 | Conclusion

All four Chatbots responded to cognitive level 1 questions better than to cognitive level 2 questions. Microsoft Copilot gave the best responses to questions, and ChatGPT 3.5 generated the best questions. None of the Chatbots generated an acceptable peer-reviewed academic reference list, and there were a few errors. Therefore, though the potential for using Chatbots for enquiry-based learning is promising but as of now these tools are not totally dependable, and learners still require human teachers for guidance and facilitation. However, the emerging sophisticated Chatbots that can be integrated into Learning Management Systems with access to academic resources hold promise for enquiry-based learning, especially if technologists collaborate with subject experts. Future research on the reliability of chatbots in sectors such as health and education, using robust methodologies, is crucial due to the high stakes involved.

Acknowledgements

Open access publishing facilitated by Auckland University of Technology, as part of the Wiley - Auckland University of Technology agreement via the Council of Australian University Librarians.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. R. Greenwood, S. Austin, K. Bacon, and S. Pike, "Enquiry Learning in the Primary Classroom: Student Teachers' Perceptions," *Education* 50, no. 3 (2022): 404–418, https://doi.org/10.1080/03004279.2020.1853788.

2. G. Kidman and N. Casinader, *Inquiry-Based Teaching and Learning Across Disciplines: Comparative Theory and Practice in Schools* (Palgrave Macmillan, 2017).

3. A. N. Riza, I. Hidaya, and P. I. Santosa, "Use of Chatbots in E-Learning Context: A Systematic Review," *IEE World AI IoT Congress (AIoT) AI IoT Congress (AIoT)*, pp. 0819–0824 (IEEE, 2023), https://doi.org/10.1109/AIIoT58121.2023.10174319.

4. L. Labadze, M. Grigolia, and L. Machaidze, "Role of AI Chatbots in Education: Systematic Literature Review," *International Journal of Educational Technology in Higher Education* 20, no. 56 (2023): 1–17, https://doi.org/10.1186/s41239-023-00426-1.

5. S. Rampf, H. Gehrig, A. Möltner, M. R. Fischer, F. Schwendicke, and K. C. Huth, "Radiographical Diagnostic Competences of Dental Students Using Various Feedback Methods and Integrating an Artificial Intelligence Application-A Randomized Clinical Trial," *European Journal of Dental Education* 28, no. 4 (2023): 925–937.

6. K. Giannakopoulus, A. Kavadella, A. Salim, et al., "Evaluation of the Performance of Generative AI Large Language Models ChatGPT, Google Bard, and Microsoft Bing Chat in Supporting Evidence-Based Dentistry: Comparative Mixed Methods Study," *Journal of Medical Internet Research* 25 (2023): e51580, https://doi.org/10.2196/51580.

7. A. A. Rahsepar, N. Tavakoli, G. H. J. Kim, C. Hassani, F. Abtin, and A. Bedayat, "How AI Responds to Common Lung Cancer Questions: ChatGPT Versus Google Bard," *Radiology* 307, no. 5 (2023): 288, https://doi.org/10.1148/radiol.230922.

8. S. Husain, S. Shahid, Z. Ansari, T. Ayoob, A. Hussain, and R. Mujahid, "To Compare the Efficiency of ChatGPT and Bard in Medical Education: An Analysis of MCQ-Based Learning and Assessment," *Annals of Abbasi Shaheed Hospital and Karachi Medical and Dental College* 29, no. 1 (2024): 56–63, https://doi.org/10.58397/ashkmdc.v29i1.782.

9. R. M. Blake and J. A. Khusid, "Artificial Intelligence for Urology Research: The Holy Grail of Data Science or Pandora's Box of Misinformation?," *Journal of Endourology* 38, no. 8 (2024): 741–747.

10. J. D. B. Featherstone and B. W. Chaffee, "The Evidence for Caries Management by Risk Assessment (CAMBRA)," *Advances in Dental Research* 29, no. 1 (2018): 9–14.

11. L. Wu, "Students' Foreign Language Learning Adaptability and Mental Health Supported by Artificial Intelligence," *Journal of Autism and Developmental Disorders* 54, no. 10 (2024): 3921–3932.

12. Y. Li, J. Xu, Y. Zhu, H. Liu, and P. Liu, "The Impact of ChatGPT on Software Engineering Education: A Quick Peek," *10th International Conference on Dependable Systems and Their Applications (DSA) DSA*

Dependable Systems and Their Applications (DSA), pp. 595–596 (IEEE, 2023).

13. D. O'Leary, "An Analysis of Three Chatbots: BlenderBot, ChatGPT and LaMDA," *Intelligent Systems in Accounting, Finance and Management* 30 (2023): 41–54, https://doi.org/10.1002/isaf.1531.

14. I. Seth, B. Lim, Y. Xie, et al., "Comparing the Efficacy of Large Language Models ChatGPT, Bard, and Bing AI in Providing Information on Rhinoplasty: An Observational Study," *Aesthetic Surgery Journal Open Forum* 15 (2023): 1–9, https://doi.org/10.1093/asjof/ojad084.

15. M. R. King, "The Future of AI in Medicine: A Perspective From a Chatbot," *Annals of Biomedical Engineering* 51 (2022): 291–295, https://doi.org/10.1007/s10439-022-03121-w.

16. M. Singh, C. Patvardhan, and C. V. Lakshmi, "Does ChatGPT Spell the End of Automatic Question Generation Research?" *IEEE International Conference on Computer Vision and Machine Intelligence (CVMI) Computer Vision and Machine Intelligence (CVMI)*, pp. 1–6 (IEEE, 2023).

17. S. Morreel, V. Verhoeven, and D. Mathysen, "Microsoft Bing Outperforms Five Other Generative Artificial Intelligence Chatbots in the Antwerp University Multiple Choice Medical License Exam," *PLOS Digital Health* 3, no. 2 (2024): e0000349, https://doi.org/10.1371/journal.pdig.0000349.

18. H. Mohammad-Rahimi, Z. H. Khoury, M. I. Alamdari, et al., "Performance of AI Chatbots on Controversial Topics in Oral Medicine, Pathology, and Radiology," *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology* 137, no. 5 (2024): 508–514, https://doi.org/10.1016/j.oooo.2024.01.015.

19. A. Gilson, C. W. Safranek, T. Huang, et al., "How Does ChatGPT Perform on the United States Medical Licensing Examination (USMLE)? The Implications of Large Language Models for Medical Education and Knowledge Assessment," *JMIR Medical Education* 9 (2023): e57594, https://doi.org/10.2196/45312.

20. J. R. Janopaul-Naylor, A. Koo, D. Qian, et al., "Physician Assessment of ChatGPT and Bing Answers to American Cancer Society's Questions to Ask About Your Cancer," *American Journal of Clinical Oncology* 47, no. 1 (2024): 17–21, https://doi.org/10.1097/COC.0000000000001050.

Appendix A

Rubric Used

	Evidence					
Criteria	**(5)**	**(4)**	**(3)**	**(2)**	**(1)**	**(0)**
Analysis, depth and accuracy	Provides accurate, extended and thorough information indicating a very high level of analysis	Provides accurate and thorough information, Indication a high analysis	Provides adequate information to the question asked indicating adequate analysis	Provides basic information with no errors indicating basic analysis	Provides basic information with no critical errors indicating some level of analysis	Provides inadequate information or information with errors indicating poor analysis
Comprehension and clarity	The information is of a high quality, in an academically acceptable style with a structure that is concise and logical	The information is in a sound academically acceptable style with a structure that is concise and logical	The information is in an academically acceptable style with a structure that is concise and logical with minor inconsistencies	The information is not in an academically acceptable style, but the structure is easy to read with minor inconsistencies	The information is not in an academically acceptable style, but the structure is not easy to read with minor inconsistencies	The information has major errors

Appendix B

List of Questions, Level 1 Questions

Oral biology	**Dental materials**	**Periodontology**	**Restorative dentistry**
What is the plexus of veins near the lateral pterygoid muscle?	What is the wavelength of the light used in dentistry to cure composite resin?	In the context of periodontology, what is an area specific curette?	In the context of restorative dentistry, what is meant by smear layer?
What type of joint is the temporomandibular joint?	Name the ingredient in the conditioner used for a dental restoration with glass ionomer cement?	In the context of periodontology, what is a PerioChip?	In the context of restorative dentistry, what are resin tags?
Which of the four muscles of mastication does not help with jaw closure?	What is the process of mixing amalgam alloy with mercury called?	In the context of periodontology, where can a Gracey 11/12 curette be used?	In the context of restorative dentistry, what are dentinal tags?
Which virus causes hand, foot and mouth disease?	What is total acid etch in restorative dentistry?	What are the parts of a periodontal curette?	In the context of restorative dentistry, what is the hybrid layer?
Which virus causes shingles?	What is the concentration of phosphoric acid etch used in restorative dentistry?	In the context of Periodontology, where can a Gracey 5/6 can be used?	What is indirect dental pulp capping?
What are enamel tufts?	Which are the metals found in dental amalgam?	What are the clinical features of necrotising gingivitis?	What is direct dental pulp capping?
What is the neonatal line in the dental enamel?	What are the components of glass ionomer cement used in restorative dentistry?	Name two types of periodontal curettes.	What are the disadvantages of air abrasion in restorative dentistry?
In the context of oral pathology, what is metastasis?	What are the contraindications for use of glass ionomer cements in dentistry?	Name two conditions that affect healing of periodontal tissues.	What are the advantages of using rubber dam for dental restorative procedures?
What are the clinical features of herpetic gingivostomatitis?	In the context of restorative dentistry, what is the difference between glass ionomer cement and resin modified glass ionomer cement?	What is oral biofilm?	What are the advantages of using the atraumatic restorative treatment technique in restorative dentistry?
In the context of oral pathology, what is papilloma?	In restorative dentistry, what is hybrid composite?	What is dental calculus?	What is the purpose of placing orthodontic separators for the placement of a stainless-steel crown using the Hall technique in restorative dentistry?

Appendix C

List of Questions, Level 2 Questions

Oral biology	Dental materials	Periodontology	Restorative dentistry
Correlate the development of the tongue to its nerve supply Describe the response of the dental pulp to deep dental caries Discuss the function of saliva in the mastication process Discuss the role of diet in causing dental caries Discuss the clinical features of oral candidiasis	Discuss the advantages and disadvantages of composite resin as a dental restorative material Discuss the advantages and disadvantages of glass ionomer cement as a dental restorative material Describe the physical and chemical properties of calcium hydroxide Describe corrosion in dental amalgam Describe the function of ferric sulphate during the pulpotomy procedure	Discuss how smoking affects periodontal health Discuss the relationship between diabetes and periodontal disease Discuss the causes of gingival redness Discuss the nonmodifiable risk factors for biofilm induced periodontal disease Discuss the causes and clinical features of peri implant mucositis and periimplantitis	Discuss the placement procedure of a stainless-steel crown on a deciduous molar using the conventional technique Explain the purpose of protecting glass ionomer cements after their placement in restorative dentistry Discuss the benefits of placing glass ionomer cements on fissures of newly erupting molars Discuss the purpose of placing resin fissure sealant materials in deep dental fissures Discuss the different methods used in dentistry to remove carious dentine

Appendix D

List of Questions—Level 3 Context-Based Questions

A 45-year-old male patient with a history of smoking presents with an asymptomatic white patch on the tongue. What could the white patch be?
A 5-year-old child has multiple cavitations in the lower and upper occlusal surfaces of the deciduous molars. The maxillary incisor teeth have cavitations mainly on the labial surfaces. What could the diagnosis be?
A 25-year-old female patient complains of recurrent blisters on her lips. She says she develops such blisters whenever she does surfing. What could the diagnosis be?
A 12-year-old child complains of pus oozing from gums in relation to a broken front tooth that is also discoloured. What could the diagnosis be?
A 6-year-old child has severe pain in a cavitated lower right second molar and the pain is keeping the child up at night. The radiograph reveals a large cavity with no periapical changes in the tooth. What could be the diagnosis?

Appendix E

Examples of Answers to Question and Average Score Allocated

Prompt	Response	Average score allocated
In restorative dentistry, what is hybrid composite?	In restorative dentistry, hybrid composite refers to a material that combines natural tooth tissue with artificial substances, offering superior mechanical, chemical and biological properties. This hybrid layer, also known as 'super dentine', is created at the adhesive interface to prevent primary and secondary caries effectively.	0 for accuracy
A 6-year-old child has severe pain in a cavitated lower right second molar and the pain is keeping the child up at night. The radiograph revels a large cavity with no periapical changes in the tooth. What could be the diagnosis?	Given the severe pain, particularly at night, and the presence of a large cavity with no periapical changes on radiograph, the likely diagnosis is irreversible pulpitis. Irreversible pulpitis occurs when the dental pulp becomes irreversibly inflamed and cannot heal on its own. Treatment such as pulpectomy or extraction may be necessary.	4.5 for accuracy

Prompt	Response	Average score allocated
Discuss the relationship between diabetes and periodontal disease.	Diabetes and periodontal disease share a complex and bidirectional relationship, meaning they significantly influence each other's progression and severity. Here's a closer look at this intricate connection: Diabetes increases risk of periodontal disease: High blood sugar: Uncontrolled blood sugar levels in diabetes create a favourable environment for the growth of harmful bacteria in the mouth. These bacteria thrive on the excess glucose present in saliva, leading to the formation of plaque and the initiation of gum inflammation. Impaired immune response: Diabetes weakens the immune system, making it harder for the body to fight off bacterial infections in the gums. This impaired defence system allows plaque bacteria to proliferate unchecked, leading to the progression of periodontal disease. Reduced blood flow: Diabetes can damage blood vessels, leading to decreased blood flow to the gums. This reduced blood supply compromises the delivery of oxygen and nutrients needed for tissue health, hindering the body's ability to repair and heal inflamed gum tissue. Periodontal disease worsens diabetes control: Increased inflammation: Chronic inflammation associated with gum disease triggers the release of inflammatory markers that can interfere with insulin sensitivity and action. This makes it harder for the body to utilise glucose effectively, leading to higher blood sugar levels and potentially worsening diabetes control. Increased stress: Periodontal disease can cause pain and discomfort, adding stress to the body. This stress response can further elevate blood sugar levels, making it more challenging to maintain glycaemic control. Consequences of the bidirectional relationship: Severity of periodontal disease: People with diabetes are more likely to develop severe forms of periodontal disease, characterised by deeper gum pockets, faster bone loss and increased tooth loss. Complications of diabetes: Uncontrolled periodontal disease can exacerbate other complications associated with diabetes, such as cardiovascular disease and kidney disease. Importance of maintaining good oral hygiene: Regular brushing, flossing and professional dental cleanings are crucial for individuals with diabetes to prevent and manage periodontal disease. Maintaining good oral hygiene helps control plaque and bacteria buildup, reducing inflammation and its impact on blood sugar control. Additional factors: Smoking: Smoking significantly increases the risk of both diabetes and periodontal disease, further compounding the negative effects on oral health. Genetics: Certain genetic factors may predispose individuals to both conditions, making them more susceptible to their development and progression. Conclusion: The relationship between diabetes and periodontal disease is a crucial aspect of overall health management. By understanding this bidirectional link, individuals with diabetes can prioritise good oral hygiene practices and regular dental care to prevent gum disease and its potential impact on their overall health. Additionally, maintaining good glycaemic control helps reduce the risk and severity of periodontal disease, creating a positive feedback loop for better oral and overall health.	3 for depth

Prompt	Response	Average score allocated
Discuss how smoking affects periodontal health.	How smoking affects periodontal health Smoking is a well-established risk factor for the initiation, progression and severity of periodontal disease. Numerous studies have demonstrated the detrimental effects of smoking on the health of the periodontal tissues. Smoking has a multifaceted impact on periodontal health: Microbial changes: Smoking alters the oral microbiome, leading to an increase in periodontal pathogens like *Porphyromonas gingivalis*. This shift in the microbial balance contributes to the development and progression of periodontal disease. Immune response: Smoking impairs the immune system's ability to respond effectively to the bacterial challenge, reducing the body's capacity to control the inflammatory response and clear the infection. This results in more severe and persistent periodontal inflammation. Tissue destruction: Smoking has been shown to increase the production of inflammatory mediators, such as cytokines and matrix metalloproteinases, which can lead to the destruction of the periodontal tissues, including the gingiva, periodontal ligament and alveolar bone. Healing impairment: Smoking can negatively impact the healing process following periodontal treatment, reducing the chances of successful outcomes. Smokers are more likely to experience recurrence of periodontal disease and tooth loss compared to nonsmokers. Dose-dependent relationship: The risk of developing periodontal disease increases with the amount and duration of smoking. Studies have found a dose-dependent relationship, where individuals with a higher number of pack-years (a measure of lifetime smoking exposure) have a higher risk of periodontal disease. Interestingly, the negative effects of smoking on periodontal health are not limited to current smokers. Former smokers also have an increased risk of periodontal disease compared to nonsmokers, although the risk is lower than that of current smokers. This suggests that the detrimental effects of smoking on the periodontium can persist even after quitting. In conclusion, smoking has a profound and multifaceted impact on periodontal health. It alters the oral microbiome, impairs the immune response, increases tissue destruction and impedes the healing process. The risk of periodontal disease is directly proportional to the amount and duration of smoking. Encouraging and supporting smokers to quit, as well as implementing targeted preventive and treatment strategies for smokers, are crucial in maintaining good periodontal health.	4 for depth