

Appendix 1: Task 0 (Baseline) questionnaire

Dear Participant,

We invite you to participate in a research study exploring the use and perceptions of AI-powered chatbots, specifically those utilizing Large Language Models (LLMs), in medical education. Your insights will help us understand how these technologies impact learning and how they can be integrated effectively into medical curricula.

Participation in this survey is voluntary. Your responses will be kept confidential and will be used solely for research purposes. The survey should take approximately 10 minutes to complete. By proceeding, you consent to participate in this study.

Thank you for your valuable time and input.

What is your KCL email address (name@kcl.ac.uk)?

Q1. What is your year group?

A100 Year 1

EMDP Year 1A

EMDP Year 1B

A100 Year 2

EMDP Year 2

Other...

Q2. Assigned ID number

Q3. What is your age?

Q4. What is your gender?

Male

Female

Prefer not to say

Other...

Q5. Are you a local or international student?

Local

International

Q6. Have you used AI LLM Chatbots to help you study? (eg clarifying content, coming up with points, explanations, writing content etc)

Yes

No

Q7. How long have you used AI LLM Chatbots?

1-3 months

4-6 months

7-9 months

10-12 months

Under 2 years

Over 2 years

Q8. How frequent do you use AI LLM Chatbots per week for studying?

1 time per week

2 times per week

3 times per week

4 times per week

5 times per week

6 times per week

7 times per week (every day)

Never/Rarely

Q9. Which AI-powered chatbots have you used? (Select all that apply)

ChatGPT

Google Gemini

Microsoft Copilot / Bing AI

Meta's Llama

Other...

Q10. Which of the following best describes how you use AI chatbots in your studies?
(Select all that apply)

Understanding complex concepts

Generating study materials (e.g., summaries, flashcards)

Assisting with problem-solving or assignments

Reviewing or editing written work

Writing essays/paragraphs

Preparing for exams

Other...

Q11. How confident are you in your ability to use AI-powered chatbots effectively?

Not confident at all

Very Confident

1 2 3 4 5

Q12. Please indicate your level of agreement with the following statements

Strongly Disagree

Disagree

Neutral

Agree

Strongly Agree

a) AI chatbots are helpful in enhancing my learning experience.

b) AI chatbots provide accurate information relevant to my studies.

c) Using AI chatbots has improved my academic performance.

- d) I find AI chatbots easy to use.
- e) I am concerned about the reliability of information provided by AI chatbots.
- f) AI chatbots encourage me to think critically about the information I receive.
- g) I prefer traditional study resources over AI chatbots.
- h) AI chatbots save me time in completing academic tasks.
- i) I am worried about potential academic integrity issues when using AI chatbots.
- j) AI chatbots will play a significant role in the future of medical education.

Q13. How do you perceive the quality of responses provided by AI chatbots?

Very Poor

Very Good

1 2 3 4 5

Q14. What limitations do you perceive in using AI-powered chatbots? (Select all that apply)

Inaccuracy of information

Difficulty prompting

Difficulty in understanding complex queries

Traditional Resources, including Google Search, are more than enough

Lack of creativity

Lack of depth in responses

User interface or formatting is confusing

Unclear Wordings

Other...

Q15. To what extent do you believe AI chatbots can replace traditional learning resources (e.g., textbooks, lectures)?

Not at all

Completely

1 2 3 4 5

Q16. After using AI LLM Chatbots, how confident do you feel in applying the information learned?

Not confident at all

Very confident

1 2 3 4 5

Q17. What potential benefits do you see in integrating AI chatbots into medical education? (Select all that apply)

Personalized learning experiences

Immediate access to information

Enhanced understanding of complex topics

Improved study efficiency

Development of critical thinking skills

Assistance with problem-solving

Support for collaborative learning

Other...

Q18. In your opinion, how important is it for medical students to be proficient in using AI tools?

Not Important

Extremely Important

1 2 3 4 5

Q19. How likely are you to continue using AI LLM Chatbots in your future studies?

Not likely at all

Very likely

1 2 3 4 5

Appendix 2: Task 1 questions

Part 1: Single Best Answer (SBA) Questions

Question 1

A 24-year-old male presents to the emergency department after a motorcycle accident. He is unable to abduct his shoulder and has significant weakness when trying to externally rotate the arm. Physical examination reveals an inability to initiate shoulder abduction and a loss of sensation over the lateral aspect of the shoulder. Which part of the brachial plexus is most likely injured?

- A. Suprascapular and axillary nerves
- B. Long thoracic nerve
- C. Medial cord
- D. Musculocutaneous nerve
- E. Radial nerve

Question 2

A 35-year-old construction worker reports weakness in his hand after a fall from a ladder. On examination, he exhibits wrist drop and weakened extension of the fingers and elbow. Sensory loss is noted on the posterior aspect of the forearm and dorsum of the hand. Which part of the brachial plexus is likely affected?

- A. Lateral cord
- B. Medial cord
- C. Posterior cord
- D. Suprascapular nerve
- E. Ulnar nerve

Question 3

A 19-year-old college athlete presents with numbness and tingling along the medial side of the forearm and hand following an incident while lifting weights. Examination reveals weakness in flexing the fourth and fifth digits and diminished hand grip strength. Which nerve of the brachial plexus is most likely compressed or damaged?

- A. Axillary nerve
- B. Long thoracic nerve

- C. Median nerve
- D. Musculocutaneous nerve
- E. Ulnar nerve

Question 4

A 32-year-old woman visits the clinic complaining of difficulty gripping objects and numbness along the lateral aspect of her palm and first three fingers. Upon examination, you notice she cannot make a fist, as the index and middle fingers remain extended while attempting to close her hand. What condition is she most likely suffering from?

- A. Axillary nerve damage
- B. Carpal tunnel syndrome
- C. Hand of benediction
- D. Radial nerve palsy
- E. Thoracic outlet syndrome

Question 5

A 40-year-old office worker presents with tingling and numbness in his thumb, index, and middle fingers, especially at night. He also reports weakness when trying to perform pinching motions, such as holding a pen or turning a key. Physical examination shows atrophy of the thenar eminence. Which of the following is the most likely diagnosis?

- A. Carpal tunnel syndrome
- B. Cubital tunnel syndrome
- C. Hand of benediction
- D. Klumpke's palsy
- E. Thoracic outlet syndrome

Question 6

Which of the following signs is most consistent with cubital tunnel syndrome?

- A. Atrophy of the hypothenar and interosseous muscles
- B. Inability to extend the wrist, resulting in a wrist drop
- C. Loss of sensation over the lateral three and a half digits
- D. Weakness in shoulder abduction and external rotation
- E. Weakness in elbow flexion and forearm supination

Question 7

A 25-year-old athlete presents with weakness when abducting the shoulder and externally rotating the arm. On examination, there is a notable flattening of the shoulder contour, and sensation is diminished over the lateral aspect of the upper arm. Which of the following clinical conditions is most likely?

- A. Erb-Duchenne palsy
- B. Klumpke's palsy
- C. Radial nerve palsy
- D. Rotator cuff tear
- E. Shoulder dislocation-associated nerve injury

Part 2: Short Answer Questions

Question 1: What are the five terminal branches of the brachial plexus?

Question 2: Which nerve of the brachial plexus is responsible for innervating the biceps brachii muscle?

Question 3: A 34-year-old cyclist falls off his bike and lands on his outstretched arm. He presents with weakness in shoulder abduction, elbow flexion, and loss of sensation over the lateral aspect of the forearm. Which part of the brachial plexus is most likely affected?

Question 4: A 50-year-old patient complains of difficulty with wrist and finger extension, resulting in a characteristic wrist drop. On examination, there is also numbness over the posterior aspect of the forearm and dorsum of the hand. Which nerve of the brachial plexus is most likely involved?

Question 5: A 27-year-old weightlifter experiences pain and weakness when trying to lift his arm above his head. On examination, you notice scapular winging when he pushes against a wall. What condition related to the brachial plexus is most likely causing these symptoms?

Question 6: A newborn is delivered via difficult labor and presents with a “claw hand” deformity affecting the wrist and fingers. The infant also exhibits weakness in the intrinsic muscles of the hand. Which condition related to the brachial plexus is most likely the diagnosis?

Question 7: A 45-year-old man presents with a history of numbness and tingling in the fourth and fifth fingers, along with weakness in finger abduction and adduction. He also has atrophy of the hypothenar muscles. What condition related to the brachial plexus could explain these findings?

Appendix 3: Task 2 questions

Part 1: Single Best Answer (SBA) Questions

Question 1:

Which nerve, originating from the lumbar plexus, is responsible for innervating the quadriceps muscle group?

- A. Femoral nerve
- B. Obturator nerve
- C. Pudendal nerve
- D. Sciatic nerve
- E. Superior gluteal nerve

Question 2 (this question is cancelled)

A 50-year-old woman presents with difficulty climbing stairs and standing up from a seated position. She also reports weakness in hip extension. Which nerve is most likely affected?

- A. Femoral nerve
- B. Inferior gluteal nerve
- C. Obturator nerve
- D. Sciatic nerve
- E. Superior gluteal nerve

Correct answer: B. Inferior gluteal nerve

Question 3:

A 45-year-old man reports pain radiating from his lower back to the posterior aspect of his thigh and down to his calf and foot. He also has difficulty with knee flexion and dorsiflexion of the foot. Which nerve is most likely implicated in this case?

- A. Femoral nerve
- B. Inferior gluteal nerve
- C. Obturator nerve
- D. Sciatic nerve
- E. Superior gluteal nerve

Question 4

A 36-year-old woman presents to the clinic with a complaint of difficulty walking, particularly when stepping out of a car. On examination, she has weakness in thigh adduction and a sensory deficit over the medial aspect of her thigh. She denies any back pain. What nerve is most likely impaired?

- A) Common peroneal nerve
- B) Femoral nerve
- C) Inferior gluteal nerve
- D) Obturator nerve
- E) Tibial nerve

Question 5

A 45-year-old woman presents to the emergency department with acute onset of severe pain in the left lower limb, which started while she was climbing stairs. On examination, she has a palpable, tender mass in the popliteal fossa, and the pain increases with passive

dorsiflexion of the foot. The patient has a history of chronic venous insufficiency. Which of the following conditions is the most likely cause of her symptoms?

- A. Achilles tendon rupture
- B. Deep vein thrombosis
- C. Femoral hernia
- D. Popliteal artery aneurysm
- E. Ruptured Baker's cyst

Question 6

Which of the following signs or symptoms is most characteristic of Charcot-Marie-Tooth disease?

- A. High-stepping gait with foot drop
- B. Claudication pain relieved by rest
- C. Glove and stocking sensory loss
- D. Severe pain and swelling after trauma
- E. Sudden, sharp back pain radiating down the leg

Question 7

A 72-year-old man presents with pain in the left hip that radiates down the lateral aspect of the thigh to the knee. He has difficulty walking and frequently loses his balance. On examination, there is weakness in hip abduction, and Trendelenburg's sign is positive on the left side. Which of the following is the most likely cause of his symptoms?

- A. Greater trochanteric pain syndrome
- B. Iliotibial band syndrome
- C. Lumbar spinal stenosis
- D. Meralgia paresthetica
- E. Osteoarthritis of the hip

Part 2: Short Answer Questions

Question 1: Which nerve originating from the lumbosacral plexus is responsible for motor innervation of the quadriceps muscle group?

Question 2: A 30-year-old man presents with weakness in hip flexion and knee extension after a motor vehicle accident. He also complains of numbness over the anterior thigh. Which nerve of the lumbosacral plexus is likely injured?

Question 3: A 50-year-old woman complains of pain radiating from her lower back to the posterior thigh and lateral aspect of the leg. On examination, she has weakness in plantarflexion and absent ankle reflex. Which nerve of the lumbosacral plexus is affected?

Question 4: A 25-year-old male athlete presents with difficulty standing on his toes and sensory loss over the sole of his foot. He denies any lower back pain. What is the most likely nerve involved, and what could be a potential cause?

Question 5: A 65-year-old woman with a history of hip replacement surgery reports weakness in hip abduction and numbness over the lateral aspect of her thigh. Which nerve of the lumbosacral plexus may have been damaged during the surgery, and what symptoms support this diagnosis?

Question 6: A 45-year-old man presents with severe pain in his lower back that radiates down the posterior thigh and into the lateral aspect of his foot. On physical examination, he has difficulty with ankle dorsiflexion and reduced sensation over the dorsum of the foot. Which nerve root is most likely compressed, and what condition is commonly associated with this presentation?

Question 7: A 60-year-old woman who recently underwent pelvic surgery presents with difficulty climbing stairs and weakness in extending her knee. She also complains of

numbness over the anterior and medial thigh. Which nerve is most likely injured, and what is a potential complication related to this nerve damage?

Appendix 4: Post Task 1 Questionnaire

QVault AI Chatbody Study - Post task 1 questionnaire

Dear Participant,

Thank you for completing Task 1 of our study. Please take a few minutes to complete this questionnaire about your experience during the task. Your responses are crucial for our research on learning methods in medical education.

This questionnaire should take approximately 5 minutes to complete. Your responses are confidential and will be used solely for research purposes.

Q1. Assigned ID number

Section 1: Instructions

Please answer the following questions based on the learning method you used during Task 1:

Group A: Used the AI Chatbot along with the provided handout.

Group B: Used traditional resources (e.g, Google search without AI features and textbook handout).

Section 2 of 5

Section A: Experience with the learning method

Q2. How easy was it to used the learning method provided during Task 1?

Very difficult

Very easy

1 2 3 4 5

Q3. How satisfied are you with the learning method you used during Task 1?

Very dissatisfied

Very satisfied

1 2 3 4 5

Q4. To what extent did the learning method help you understand the topic (brachial plexus) during task 1?

Not at all

A great deal

1 2 3 4 5

Q5. How efficient was the learning method in helping you find the information you needed?

Very Inefficient

Very efficient

1 2 3 4 5

Q6. After using the assigned learning method, how confident do you feel in applying the information learned?

Not confident at all

Very confident

1 2 3 4 5

Section 3 of 5

Section B: Perceptions of the learning method

Q7. How would you rate the quality of the information provided by your assigned learning method?

Very poor

Excellent

1 2 3 4 5

Q8. To what extent do you believe the information obtained from your given learning method was accurate?

Not accurate at all

Completely Accurate

1 2 3 4 5

Q9. How would you rate the depth of content provided by your given learning method?

Very superficial

Very in-depth

1 2 3 4 5

Q10. How easy was it to understand the information provided by your given learning method?

Very difficult

Very easy

1 2 3 4 5

Q11. How engaging was the learning method in maintaining your interest during the task?

Not engaging at all

Extremely engaging

1 2 3 4 5

Section 4 of 5

Section C: Comparison with previous learning experiences

Q12. Compared to your usual study methods, how did the learning method you used during Task 1 perform?

Much worse

Much better

1 2 3 4 5

Q13. How do you feel the learning method affected your ability to think critically about the subject matter?

Strongly hindered

Strongly aided

1 2 3 4 5

Q14. How likely are you to use this type of learning method in your future studies?

Very Unlikely

Very likely

1 2 3 4 5

Section 5 of 5

Section D: Perceptions of AI Chatbots

Q14. What is your overall attitude toward the use of AI Chatbots in medical education?

Very negative

Very positive

1 2 3 4 5

Q15. How useful do you believe AI chatbots are in supporting medical education?

Not useful at all

Extremely useful

1 2 3 4 5

Q16. Do you have any concerns about using AI chatbots in your studies (Select all that apply)

Accuracy of information

Over reliance on technology

Ethical considerations

Privacy issues

Impact on critical thinking skills

Lack of human interaction

No concerns

Other...

Q17. How interested are you in using AI chatbots in your future studies?

Not interested at all

Extremely Interested

1

2

3

4

5

Appendix 5: Post Task 2 Questionnaire

QVault AI Chatbody Study - Post task 2 questionnaire

Dear Participant,

Thank you for completing Task 2 of our study. Please take a few minutes to complete this questionnaire about your experience during the task. Your responses are crucial for our research on learning methods in medical education.

This questionnaire should take approximately 5 minutes to complete. Your responses are confidential and will be used solely for research purposes.

Assigned ID number

Section 1: Instructions

Please answer the following questions based on the learning method you used during Task 2:

Group A: Used traditional resources (e.g, Google search without AI features and textbook handout).

Group B: Used the AI Chatbot along with the provided handout.

Section 2 of 6

Section A: Experience with the learning method

Q1. How easy was it to use the learning method provided during Task 2?

Very difficult

Very easy

1 2 3 4 5

Q2. How satisfied are you with the learning method you used during Task 2?

Very dissatisfied

Very satisfied

1 2 3 4 5

Q3. To what extent did the learning method help you understand the topic (lumbosacral plexus) during task 2?

Not at all

A great deal

1 2 3 4 5

Q4. How efficient was the learning method in helping you find the information you needed?

Very Inefficient

Very efficient

1 2 3 4 5

Q5. After using the assigned learning method, how confident do you feel in your understanding of the lumbosacral plexus?

Not confident at all

Very confident

1 2 3 4 5

Section 3 of 6

Section B: Perceptions of the learning method

Q6. How would you rate the quality of the information provided by your assigned learning method?

Very poor

Excellent

1 2 3 4 5

Q7. To what extent do you believe the information obtained from your given learning method was accurate?

Not accurate at all

Completely Accurate

1 2 3 4 5

Q8. How would you rate the depth of content provided by your given learning method?

Very superficial

Very in-depth

1 2 3 4 5

Q9. How easy was it to understand the information provided by your given learning method?

Very difficult

Very easy

1 2 3 4 5

Q10. How engaging was the learning method in maintaining your interest during the task?

Not engaging at all

Extremely engaging

1 2 3 4 5

Section 4 of 6

Section C: Comparison with previous learning experiences

Q11. Compared to your usual study methods, how did the learning method you used during Task 2 perform?

Much worse

Much better

1 2 3 4 5

Q12. How do you feel the learning method affected your ability to think critically about the subject matter?

Strongly hindered

Strongly aided

1 2 3 4 5

Q13. How likely are you to use this type of learning method in your future studies?

Very Unlikely

Very likely

1 2 3 4 5

Section 5 of 6

Section D: Overall Perceptions After Using Both Learning Methods

Q14. After using both learning methods (AI Chatbot and Traditional resources), which do you prefer?

Strongly prefer traditional resources Strongly prefer AI chatbot

1 2 3 4 5

Q15. Which learning method do you feel improved your understanding of the topics more effectively?

Traditional Resources

AI chatbot

Both equally

Neither

Q16. Do you anticipate changing your study habits based on your experiences with these learning methods?

Not at all

Extremely

1 2 3 4 5

Q17. How likely are you to recommend the use of AI chatbots to your peers for studying medical topics?

Very unlikely

Very likely

1 2 3 4 5

Section 6 of 6

Section E: Perceptions of AI Chatbots

Q18. What is your overall attitude towards the use of AI chatbots in medical education after completing both tasks?

Very negative

Very positive

1 2 3 4 5

Q19. How useful do you believe AI chatbots are in supporting medical education after your experiences in this study?

Not useful at all

Extremely useful

1 2 3 4 5

Q20. How interested are you in using AI chatbots in your future studies after participating in this study?

Not interested at all

Extremely interested

1

2

3

4

5

Q21. Do you have any concerns about using AI chatbots in your studies after your experiences in this study? (Select all that apply)

Accuracy of information

Over-reliance on technology

Ethical considerations

Privacy issues

Impact on critical thinking skills

Lack of human interaction

No concerns

Other...

Appendix 6: Focus Group Transcripts

Day 1 Focus Group 1 (20th Nov)

1. Describe your experiences using chatbot and traditional experiences

Took a lot longer to filter through information in traditional.

Chatbot was faster

Traditional gave you everything

Enjoyed a mix of both

Speed - depended on the type of question - conditions you didn't know - then traditional would have been quicker. But if it was single answer then chatbot was

The aim was to get the answer down, not to learn the information, he won't remember the information at all. Main goal was

Chatbot was useful because it could provide and double check the answer, to verify it was helpful. When they didn't have the chatbot he had to use the traditional sources and notes, Chatbot took some time to respond, some of the time was spent waiting for the chatbot to respond.

Critical thinking questions - googling and using notes enhanced the critical thinking instead of in the chatbot. Chatbot didn't encourage too much critical thinking. Very med student friendly as it gave mnemonics.

Were there any surprises?

- Character limit
- The mnemonic function - was a good thing but it would be even better if it had a toggle to switch off the mnemonic function (as not everything needs mnemonics) - Sometimes the reply didn't really work
- Cannot send another question while the chatbot is thinking

2. How did your thoughts evolve about using

(3) Overall quite positive, not generalised like chat gpt and it was customisable which was good, sometimes gives too much information which I don't necessarily need for the lecture. Don't know if the information is legit using general chatGPT, but this one seems more legit.

(2) Was surprised to be able to use a chatbot for reputable information within the first

few years of university, instead of using textbooks.

(4) You usually have to spend a lot of time refining your prompts for chatGPT, but Lenny gives you the relevant med school related information as the prompt is already refined.

3. Question 3

(1) Traditional - you have to go through all of the information, and the residual information is quite useful for understanding, whereas with a chatbot you can ask the specific question. Chatbot is better for specific questions but traditional

(2) Not sure how you would use the chatbot to understand a whole topic, feels like chatbot is good for filling in the gaps in the lecture knowledge.

4. How did each learning method impact understanding and retention (2) The nature of the exercise made it so that he would just get the answer, and not enough time to actually understand the information. This method didn't allow retaining the information.

(3) The nature of the tasks today - you have time constraints so it was a bit hard to concentrate on learning.

And also anatomy is more about memorising, not understanding.

(4) If the purpose was learning over answering questions, then we would have learned a bit more. We didn't know if the answers we got were correct at all, and because of the time con

Did you try to get the right answer or were you just trying to fill in the boxes?

Both, mostly because of the time constraint but also wanted to get the questions right.

5. How would you evaluate... usability etc.

(4) Naturally engagement with traditional methods would allow more learning. There is merit in the process of finding out the information for yourself and learning things in the process, as you are being curious about the topic and engaging with the resources more. However Lenny is good at

(1) Also need to find reputable sources through traditional sources, otherwise you're not going to get the right answer. By sifting through the materials by yourself and diving into other topics at the same time you learn more.

(3) Lenny is more question and answer, question and answer, so there may not be as much understanding.

6. Any specific features?

(3) Mnemonics - see that could be good in memorising, other features of qVault could be very useful as well, e.g. podcasts, anki generation etc.

7. Challenges or limitations?

(4) lack of critical thinking with the chatbot - a little bit of surrounding information would be good. If she got another random question on the same topic, she wouldn't know how to answer the questions. (Isaac says this is more of a resource constraint but acknowledges this point). But a good thing is that you don't have to refine the prompt in order to get a medical school level answer.

(2) could be useful to show diagrams - the chatbot didn't have diagrams. He feels like the delay is okay for a while, but if he were to use it for a long time it could be frustrating to use. One feature to improve on is the scrolling aspect - it should scroll to the bottom (didn't happen for two people)

(3) Interface - the follow up between the questions is good in chatGPT, but maybe not in Lenny. Even when you used the reply function, the responses were not specific to the follow up question. Also there is a delay. Would prefer instantaneous responses and

8. How did your experience with it influence your view?

(4) More inclined to use Lenny if it is more refined in the future, if it deals with the delays and also delves into a little bit more surrounding it. But is happy about the specificity of Lenny and how it can refine to medical school standard

(2) in its current state would only use it from time to time, but will look out for similar models which would be useful for exams. Would be useful to go through the missing parts of exams just to fill in the knowledge missing from lecture slides or clarify things that they are unsure about.

(1) Previously would not have been inclined to use ChatGPT or AI related sources

9. Suggestions for improvement

- Anki function
- Fix delay if possible
- Mnemonic toggle

- Diagrams or pictures
- Maybe reputable sources for the trust aspect
- Question generation

10. What roles do you envision - how could they enhance existing learning methods?

(2) The ideal scenario is to upload the powerpoint or transcript, and it could generate anki cards and questions, and include diagrams - so you can do spaced repetition and also test your knowledge. Also test your own knowledge over time and see progress.

(4) Being able to generate a whole booklet of questions would be really useful.

(3) good quality questions - maybe collaborate with the universities/med schools to make sure the content is more accurate.

11. Anything you want to add?

(4) The UI is very clean, easy to use, and despite the delays it was very good to use. Bearable for the purposes of this task.

(1) King's specific marketing is also good

(2 and 3) - I like Lenny.

Day 1 Focus Group 2 (20 Nov)

Key Themes and Insights

1. Experiences Using Chatbot and Traditional Resources

Chatbot give concise responses vs Google searching gives complicated info (3)

Hard to search directly for the answer in Google, but in chatbot it gives a direct answer (3)

AI simplifies process to gain info, but need to learn how to use AI too, or else not helpful in knowledge retention (1)

Traditional: all material given, can see what is related

Chatbot needs user input and cannot see how everything is related, more useful when specific query in mind instead of learning entire topic

Chatbots particularly useful in diagnosis questions vs google

Need to keep chatbot content accurate (4)

Best to train it to be tailored to curriculum – ensure relevance (4)

More useful if add reference in chatbot responses - credibility (everyone agrees)

2. Speed and Retention

Latency – discourages users, may just sway people back to Google search (1,4)

Waiting time / notifications/ technical issues may disrupt learning

3. AI Handouts

Handouts have distinct details, much quicker to get details than chatbot (only direct factual questions, chatbot better in diagnosis questions) (4)

Liked how the handout give only relevant info (4)

AI-generated handout vs textbook: Handouts for basic knowledge, use textbooks for extra info (1)

4. Evolution of Thoughts on Chatbots

Little chatbot exposure before, only the AI overview in Chrome, more open to using chatbot after this (2)

Will only use chatbots when confused with medical concepts instead of learning new concepts (3)

5. Learning and Retention

Understand lecture slides: use anki (4)

AI for understanding harder topics (4)

Only use chatbot when confused with something (2)

More useful only in consolidating already learnt basic knowledge (3)

Useful in learning & understanding concepts, not necessarily lead to better exam results (4)

Chatbots may make students lazy in learning, requires discipline to memorize content (2,3)

6. Specific Features and Challenges

Mnemonics (only) sometimes useful (1,3)

interested in the AI question generation feature - Better than passmed since they are personalised to uploaded lectures

Add diagrams, Anki generation, and question creation.

7. Roles in Learning Methods

Still trust question banks more than AI-generated questions (1)

Will trust AI more if it is trained based on past papers (all agree) – since users do not have knowledge whether all the content is true

8. Others

Best to make it free to be accessible to everyone (1) (discussed cost issues)

Day 2 Focus Group 3 (11 Dec)

1. Experiences with Learning Methods

Can you describe your experiences using both the AI chatbot and traditional resources during the tasks?

Probe: What stood out to you about each method? Were there any surprises?

001:

- More engaging than safari or google and more personalised
- AI responds to exactly what you ask, but google will give a general response to similarly asked questions

003:

- So much info in google/ safari, so in AI, it is more concise and everything is in one thing
- Google was giving loads of information, some also contradictory -> eg giving 2 separate answers

- Makes it easier to learn a list of conditions
- Less helpful to learn one specific thing/concept in detail, but to learn eg many symptoms is better (memorisation easier to learn using AI than understanding)

004:

- Do not like mnemonic, as it enforces brute memorisation rather than understanding

005:

- AI gives mnemonics to help learn and remember, and finds it useful

2. Perception Changes

How did your perceptions of AI chatbots evolve, if at all, from before the study to after using them during the tasks?

Probe: What specific experiences influenced your opinions positively or negatively?

001:

- Agree with 005, same points

003:

- Started using GPT whilst using passmed, to make the passmed explanations simpler and shorter
- Worried about inaccuracies, but if AI is specific enough, then answers seem to be more accurate

004:

- AI more tailored than google

005:

- Trust qvault more than chatgpt as it is tailored towards medical knowledge as chatgpt is more general

3. Comparative Effectiveness

In what ways did the AI chatbot and traditional resources differ in helping you understand medical topics?

Probe: Can you give examples of when one method worked better than the other?

001:

003:

- Google is frustrating when it gives contradictory answers, so she uses AI to confirm the case

004:

- Nice introduction to a topic as you can read a concise description
- Can double check with using AI, as you can ask "are you sure this is correct?"

005:

- Chatbot for learning first time easier as it can give an overview before being swarmed with information

4. Impact on Learning

How did each learning method impact your understanding and retention of the material?

Probe: Did one method encourage deeper learning or critical thinking? Why?

(Moderator's Note: Probe carefully to distinguish between perceived and actual retention.)

001:

- QV is a bit longer than GPT, but it is similar and not long enough to discourage from using QV

003:

- Used mostly passmed, but also some youtube
- Gpt can also exclude additional details not needed
- Takes longer than GPT to generate but answers are better and more relevant

004:

- QV Slower than GPT
- Don't like mnemonics

005:

- Use osmosis, lectures and anki
- QV slower than GPT
- Would be better if there was an upload feature on QV

5. Usability and Engagement

How would you evaluate the usability and engagement levels of the AI chatbot compared to traditional resources?

Probe: Were there any specific features that enhanced or detracted from your learning experience?

(Moderator's Note: Encourage discussion of both usability and engagement to balance the responses.)

001:

- QV gives more information so makes life easier

003:

- QV is really good, but the problem is even with rephrasing questions, she gets the same response so it is annoying

004:

- Finds more useful than google
- More concise than google

005:

- Also finds more useful than google

6. Challenges Encountered

What challenges or limitations did you face when using the AI chatbot and traditional resources?

Probe: How did you navigate or overcome these challenges? Did either method create unique difficulties?

001:

- Prompt wrong, so takes more time as he needs to generate a new question so cannot finish
- Agree with bullet point 3 of 003

003:

- Cannot finish for both, as QV takes too long to generate
- Keep questions as short as possible to generate questions faster
- If unlimited time, can retain more information as information is the same as it is all tailored towards medicine so answers vary less if irrelevant

004:

- Learnt how to use QV better so could improve with timing

005:

- Not enough time to finish both, but feels with GPT she would have finished as it is faster
- Word limit so could not ask whole question, so had to segment up one question into smaller bits
- Feels that with time pressure, she learnt not much, as it was just regurgitation (no understanding or memory gained)

7. Influence on Future Study Habits

How might your experiences with AI chatbots influence your future study habits or strategies?

Probe: Do you see yourself integrating AI tools into your routine, and if so, how?

001:

- Will use more in future

003:

- Very keen to use QV more in future and "would recommend to a friend"
- Would use to learn content and also revise

004:

- Will use more for revision

005:

- Will use more for revision

8. Role of AI in Medical Education

What potential roles do you envision for AI chatbots in medical education?

Probe: How could they complement or enhance existing learning methods?

001:

- Does not know much about OSCE so did not comment much

003:

- Thinks QV could help with OSCE with remembering guidelines or policies but less so with the communications skills or the human-human interaction

004:

- Talking to real people is better than AI for OSCE practise as it is hard to simulate

005:

- AI/QV could generate scenarios for OSCE practise
- If it generates a mark scheme on the scenario, it could be used to practise the content

9. Suggestions for Improvement (*Optional: Ask if time allows.*)

What improvements would you suggest for integrating AI chatbots into medical education to better support learning?

Probe: Are there specific features or functionalities you think would make them more effective?

001:

- Faster will be better
- Add image interpretation function

003:

- Faster will be better
- UI and layout is very good -> modern and sleek -> easy to use

004:

- Faster will be better
- Give citations and references will be better

005:

- Faster will be better
- Increase word count as it is a little bit restricting (eg 500 words)
- Give citations and references will be better

Appendix 7: Focus Group Question Set

1. Experiences with Learning Methods

- Can you describe your experiences using both the AI chatbot and traditional resources during the tasks?

Probe: What stood out to you about each method? Were there any surprises?

2. Perception Changes

- How did your perceptions of AI chatbots evolve, if at all, from before the study to after using them during the tasks?

Probe: What specific experiences influenced your opinions positively or negatively?

3. Comparative Effectiveness

- In what ways did the AI chatbot and traditional resources differ in helping you understand medical topics?

Probe: Can you give examples of when one method worked better than the other?

4. Impact on Learning

- How did each learning method impact your understanding and retention of the material?

Probe: Did one method encourage deeper learning or critical thinking? Why?

(Moderator's Note: Probe carefully to distinguish between perceived and actual retention.)

5. Usability and Engagement

- How would you evaluate the usability and engagement levels of the AI chatbot compared to traditional resources?

Probe: Were there any specific features that enhanced or detracted from your learning experience?

(Moderator's Note: Encourage discussion of both usability and engagement to balance the responses.)

6. Challenges Encountered

- What challenges or limitations did you face when using the AI chatbot and traditional resources?

Probe: How did you navigate or overcome these challenges? Did either method create unique difficulties?

7. Influence on Future Study Habits

- How might your experiences with AI chatbots influence your future study habits or strategies?

Probe: Do you see yourself integrating AI tools into your routine, and if so, how?

8. Role of AI in Medical Education

- What potential roles do you envision for AI chatbots in medical education?

Probe: How could they complement or enhance existing learning methods?

9. Suggestions for Improvement (*Optional: Ask if time allows.*)

- What improvements would you suggest for integrating AI chatbots into medical education to better support learning?

Probe: Are there specific features or functionalities you think would make them more effective?