

11 Key Strategies in Using AI Chat GPT to Develop HOTS-Based Entrepreneurship Questionnaires

Almaida Vebibina ^{1,2,*}, Rusijono ¹, Andi Mariono ¹, Ahsan Muafa ¹, Dydik Kurniawan ¹, Ragil Sugeng Dewantoro ¹, Dian Noer Asyari ¹, Sussi Widiastuti ¹, and Jarot Suseno ¹

¹ Department of Educational technology, Faculty of Education, Universitas Negeri Surabaya, Surabaya 60213, Indonesia;

² Department of Cosmetology Education, Faculty of Engineering, Universitas Negeri Medan, Medan 20222, Indonesia.

* **Corresponding author:** almaidavebibina@unimed.ac.id.

ABSTRACT: Developing entrepreneurship assessment questions that measure higher-order thinking skills (HOTS) is a challenge that requires time and in-depth analysis. Artificial intelligence such as Chat GPT is here to simplify the process. However, its use in creating HOTS-based questions in entrepreneurship still faces obstacles. This research is proposed based on the background of various obstacles in the use of Chat GPT to develop an effective questionnaire that meets the learning objectives. To bridge the gap, this study explores the potential of Chat GPT in crafting assessment questions that focus on higher-order thinking by utilizing existing strategies to enhance critical, analytical, creative and problem-solving thinking in entrepreneurship education. A total of 20 entrepreneurship experts at Surabaya State University were involved in this research. We developed LLM (Large Language Model) using the approach to create HOTS-based multiple-choice questions (MCQs) from an advanced course context. Subsequently, we evaluated 470 MCQs generated by Chat GPT-4 and 337 MCQs generated by Chat GPT-3.5 as well as 595 MCQs corresponding to higher-order cognitive levels. An assessment instrument was used to evaluate the quality of questions generated through 23 items adopted from previous studies. To test the difference in performance, the Mann-Whitney Test was analysed with a p value < 0.05. The results showed that ChatGPT-4 was able to generate questions with clear language usage, providing a single correct answer, as well as a high-quality selection of distractors. In addition, the questions developed by ChatGPT-4 are in accordance with HOTS aspects. This finding can be utilized by teachers as the latest generative model to increase the bank of diverse multiple-choice questions without replacing human expertise.

Keywords: generative AI, Chat GPT, bloom's taxonomy, education technology, instructional design, entrepreneurship.

I. INTRODUCTION

Entrepreneurship assessment questions are constantly evolving to match the changing market demands and social dynamics. The established assessment system with proposed evaluation tools using HOTS (Higher Order Thinking Skills) encourages the development of higher order thinking skills to keep up with the skills of the century [1, 2]. The 21st century skills partnership highlights that competencies in critical thinking, problem solving, and creativity are essential skills needed to achieve success in today's learning and working world [3]. However, creating an entrepreneurship assessment tool that can measure higher order thinking skills is not easy and fast because teachers need a good level of analytical skills and understanding of the context of HOTS questions. Some researchers have found that teachers struggle to develop assessment questions that measure analysis, synthesis and evaluation skills [4, 5].

Meanwhile, large language models based on artificial intelligence in various applications such as bard, quiz and Chat GPT have widely influenced society including in education [6, 7]. Among the latest moderation tools,

Chat GPT have been successful in the business graduation exam at Minnesota Law School by answering multiple choice assessment questions and essay questions [8]. Similarly, [9] found that Chat GPT performed well in process analysis questions based on case studies of MBA courses specifically operations management. Furthermore, [10] revealed Chat GPT can pass the main accounting certification exam through additional courses and training. In line with these studies, [11] investigated further in the field of financial accounting. They found that from Chat GPT answers using multiple-choice question banks can provide accurate answers to accounting number-based assessment questions.

In addition, undergraduate students of entrepreneurship education in the Southeast Region of the United States thought Chat GPT helped them become more creative and come up with new ideas, although some students stated that the use of Chat GPT did not increase creativity. Despite the divergent views, the study claimed Chat GPT can increase effectiveness in performing repetitive tasks thus reducing workload or streamlining work in entrepreneurship education [12]. In line with these findings, [12] described how AI (Artificial Intelligence) technology can play an important role in idea generation, evaluation, and decision-making in various types of businesses. LLM (Large Language Model) has broad and deep insights, which can be used for theory development in the field of entrepreneurship. In the marketing domain, AI makes a significant contribution to increasing sales by offering digital marketing strategies [12].

Despite the many scientific discourses adopting artificial intelligence and its implications [13-15], there is still a gap in empirical research examining the potential of Chat GPT in entrepreneurship education to create HOTS-based evaluation questions. Therefore, the researcher sought to bridge the challenge by exploring Artificial Intelligence using the Chat GPT model to fulfill a significant research gap. The focus of this study is to create multiple choice questions (MCQs) on business introduction, marketing strategy, product development, promotion strategy, business management. The purpose of this research is to explore the potential of Chat GPT through strategies developed to create assessment questions that can encourage critical, analytical, problem-solving, and creative thinking skills as part of the abilities that are needed in the current era. We created questions that represent actual evaluation questions so that they serve as a substitute for the evaluation questions themselves.

The use of GPT Chat can be utilized for the creation of entrepreneurship questions by presenting practical examples, case studies, extensive data, as well as saving time and effort in the creation of evaluation questions through the presented materials. Instead of wasting time, teachers can save time in drafting various entrepreneurship questions by emphasizing accuracy and content relevance. Teachers can also gain experience and increase knowledge on assessment questions that support analysis, synthesis and synthesis thinking through the use of Artificial Intelligence. We believe that these 11 approaches can be used to develop diverse questions that focus on higher-order thinking. These strategies are designed to empower teachers to create more meaningful assessment materials, improve learning effectiveness, and ultimately, help students better master marketing concepts. To investigate how Chat GPT can create high-quality HOTS-based assessment questions in entrepreneurship courses, we analyzed the following questions:

RQ1: To what extent does Chat GPT generate quality MCQs to meet the requirements of entrepreneurship assessment questions?

1. Does it provide sufficient information with clear language?
2. Does it provide correct answers along with quality exemptions?
3. Does the prompt code used conform to correct syntax and logic?

RQ2: What is the MCQ's level of conformity with higher order cognitive levels?

II. LITERATURE REVIEW

1. AI IN ENTREPRENEURSHIP EDUCATION

In recent years, there has been increasing interest in the use of artificial intelligence especially in entrepreneurship education [12]. There are many potentials of artificial intelligence that can be utilized for entrepreneurship education [13, 14]. Among other artificial intelligence tools, Chat GPT is the most frequently studied in research in the field of entrepreneurship. Chat GPT is an application designed by open AI using

software based on the Generative Pre-trained Transformer (GPT) architecture, which has then been enhanced through reinforcement learning techniques with human feedback [16].

According to [17], the main success of the Chat GPT application is based on its ability to perform large-scale pre-training by utilizing information from diverse sources on the internet, so that the model can learn a very large amount of data. Adapting these advantages allows Chat GPT to serve a variety of specific learning content and complex assessment questions [18]. This proves that LLM is able to process and understand various relevant contexts in various fields and types of assessment [19], for example, the use of Chat GPT to optimize the learning process of entrepreneurship through providing extensive knowledge, comprehensive analysis and assisting users in the formation of knowledge about entrepreneurship [20]. Thus, Chat GPT not only helps provide information, but allows users to engage in interactive simulations and entrepreneurial exercises using the model.

Because Chat GPT uses natural language processing, which can interpret and answer questions directly, it allows teachers to more easily organize learning materials, including multiple versions of exams, learning outcomes assessments, syllabi, and rubrics for entrepreneurship courses [18]. Chat GPT's capabilities make it easy for teachers to be more efficient in designing and managing the learning process. The use of artificial intelligence can also help students in various ways [21]. AI systems have facilities such as easy access to information that is more structured and directed than conventional search robots. With ChatGPT's ability to respond to instructions given, this system can be utilized for exploring and developing entrepreneurial ideas [22]. Students can also test their skills before exams and get help to complete entrepreneurial tasks [23]. In addition, Chatbot applications such as Chat GPT, can help students to identify business opportunities more effectively [24]. They can utilize this tool to analyze market trends, evaluate consumer needs, and design more targeted business strategies.

Previous research has discussed how Chat GPT is integrated to generate entrepreneurial products [25], and help market products [25]. This has a huge impact on economic organizations, as more and more are realizing the important role of AI technology in supporting business processes. In addition, artificial intelligence is considered as a driving factor for entrepreneurial development that can help in shaping new innovations with different methods or approaches [26]. By increasing the use of ChatGPT and other AI technologies, entrepreneurs are ready to embrace and implement innovations and capitalize on new trends emerging in the market [27]. This shows their readiness to adapt and take advantage of technological developments to support business growth.

While the capabilities of Chat GPT are impressive, there are challenges, including the accuracy of responses to more specific entrepreneurial contexts. In addition, business disciplines often involve complex data concepts and scenarios, which pose challenges for Chat GPT in computation especially in more advanced process analysis questions [9]. Therefore, it is imperative to investigate Chat GPT's ability to address entrepreneurial problems and its potential to assist teachers in entrepreneurship evaluation especially in the context of developing questions that test not only theoretical knowledge, but also problem-solving and innovation skills in real business situations.

2. *CHAT GPT AND HOTS-BASED EVALUATION ASSESMENT*

The GPT Chat model has gone through a process of learning and refinement since it was first launched in 2018. In its development, Chat GPT has been trained using a variety of data sets, including books, articles and web pages. With these enhancements, Chat GPT is able to interact in conversations in a human-like manner, answering questions, generating sophisticated text and creative content, making it difficult to distinguish from responses given by humans [28].

The majority of previous research has explored the knowledge of Chat GPT's ability to assess entrepreneurship education questions [11, 29]. However, utilizing the potential of Chat GPT to develop Higher Order Thinking Skills (HOTS)-based assessment questions, especially entrepreneurship courses, is still very limited, which provides an opportunity for researchers and practitioners to improve Chat GPT's ability to create HOTS-based questions. This reason encourages researchers to explore Chat GPT through various approaches in the context of entrepreneurship. Previously, users tended to prefer asking questions or giving simple instructions rather than composing complex texts that refer to higher-order problem solving. This imbalance is

the main motivation in the development of Chat GPT instructions for the creation of HOTS-based assessment questions.

(HOTS) measurement focuses on creating questions with the measurement process through the HOTS aspect assessment model [30]. The application of Bloom's Taxonomy (analysis, synthesis and evaluation) as an effort to define HOTS in the preparation of HOTS-based entrepreneurship test questions. The concept of HOTS questions contains elements of critical thinking and problem solving [31]. Thinking about high-level intelligence has a high level of difficulty and a long time, because it involves processing information, this activity includes the ability to gain new knowledge from available data. With an emphasis on these factors, the researcher considered using Chat GPT in the creation of HOTS-based questions. The potential and ease with which artificial intelligence technology can quickly generate text based on given instructions presents a significant opportunity for teachers to create innovative and relevant HOTS questions, and improve efficiency in the assessment process.

Scholars in previous studies have revealed the use of AI can speed up teaching, assessment and other administrative tasks [9]. As LLM uses complex algorithms and statistical models, it makes it possible to understand and learn intellectual tasks just like humans [23]. In fact, generative AI is currently capable of generating new content such as text, images, program code. With the potential of ChatGPT, this technology can be relied upon to generate text answers for high-level cognitive tasks, thus enabling the improvement of learning quality especially in the creation of evaluation questions that measure higher-order thinking skills.

III. METHOD

1. DATA SET

AI-based methods are widely used to explore their potential in entrepreneurial [32, 33]. One of them is to develop HOTS-based MCQs. In this research, the AI used are Chat GPT-4 and Chat GPT-3.5. Both AIs were chosen because they can produce questions consistently with the same standard, without being influenced by personal bias. The 11 strategies were evaluated by 2 professors and 2 entrepreneurship experts at Universitas Negeri Surabaya who have experience developing HOTS questions. Prompts were created based on this approach and run repeatedly on Chat GPT to improve its performance.

To compare the performance of the two models using existing strategies, a survey was conducted by 20 experts who work as teachers of entrepreneurship either senior or junior at Universitas Negeri Surabaya. The expert is an entrepreneurship teacher in a public school who is an expert in HOTS question preparation and the use of Chat GPT. We used a rubric to assess the quality of questions generated by Chat GPT-4 and Chat GPT-3 through 23 items that have been provided by previous research on the linkage of Chat GPT with high-level cognitive levels with categories on clarity, relevant, grammar, prompt, goal achievement [34-37]. The entrepreneurship material developed refers to the book [38].

For each prompt version, we analysed correctly generated MCQs versus incorrectly or ambiguously generated MCQs. Furthermore, each prompt approach was measured for validity using a scale of 1-5. Reverse coding was performed for negative adjustment questions. The rater agreement test (Krippendorff Alpha Test) is used to evaluate the level of consistency between raters with α values ranging from 0 to 1. Analysis of differences in Chat GPT performance was conducted using the Mann-Whitney U test. Results were considered statistically significant if $p < 0.05$, unless otherwise indicated.

2. CLASSIFICATION OF HOTS LEVELS-BLOOM TAXONOMY

In the text, the information processing answers from Chat GPT were categorized as HOTS (Higher Order Thinking Skills) levels in Bloom's Taxonomy [39]. The following table 2 shows the levels of higher cognitive learning for MCQ development [40].

Table 1. Bloom's taxonomy Coding Scheme High Cognitive Levels.

Cognitive Level	Definition	Keywords
Analyzing	Decompose information or concepts into parts and understand how the parts relate to each other.	Identifying, comparing, connecting, contrast, diagram, distinguish, illustrate
Synthesis	Bringing together different elements, ideas or information to create a new concept or theory	Compile, design, combine, develop, integrate, construct
Evaluating	Judging based on certain criteria or standards about the material	Assess, conclude, criticize, evaluate, determine, examine

3. QUESTION TYPE

In this study, we aligned the HOTS levels into entrepreneurship question topics. We mapped the 5 materials selected to be developed at the HOTS level to the question types to be created as follows:

- Business Introduction (BI) - Basic knowledge to understand business operations, business environment and managerial skills to run a business.
- Marketing Strategy (MS) - A plan designed to achieve marketing objectives that involves analyzing the market, customer needs, and competition as well as promotion, increasing sales.
- Product Development (PD) - Activities related to the design, development and marketing of new and existing products.
- Promotion Strategy (PS) - A plan designed to introduce, promote, and sell a product or service to a target market. This strategy aims to increase brand awareness, sales, and customer loyalty.
- Business Management (BM) - Activities that involve planning, organizing, directing, and controlling resources to achieve set goals.

Classifying these question types makes it possible to target higher-order cognitive processes. This is done to ensure that the right question types are targeted according to the learning objectives.

4. USER INPUT

This multiple-choice question (MCQ) creation workflow requires the user to provide information about the entrepreneurial theme of the book. In addition, the user to provide the title, a brief description of the course, a list of modules taught, HOTS concepts and strategies that have been summarized. This information is used to determine the subject matter designed and the learning objectives to be achieved. Next, the user applies the topic to the LLM by referring to the strategies provided. Some prompts were accompanied by instructions for use, while others provided moorings of prompt usage information. LLM-generated questions are compared with the material in the module as parameters.

5. MCQ GENERATION

To construct the MCQs, we provided information on the entrepreneurship course as well as targeting higher order cognitive levels (HOTS). Then, internal resources for MCQ design, along with input from users, were incorporated into the prompts submitted to Chat GPT. Figure 1 in detail outlines each step involved in the MCQ generation process.

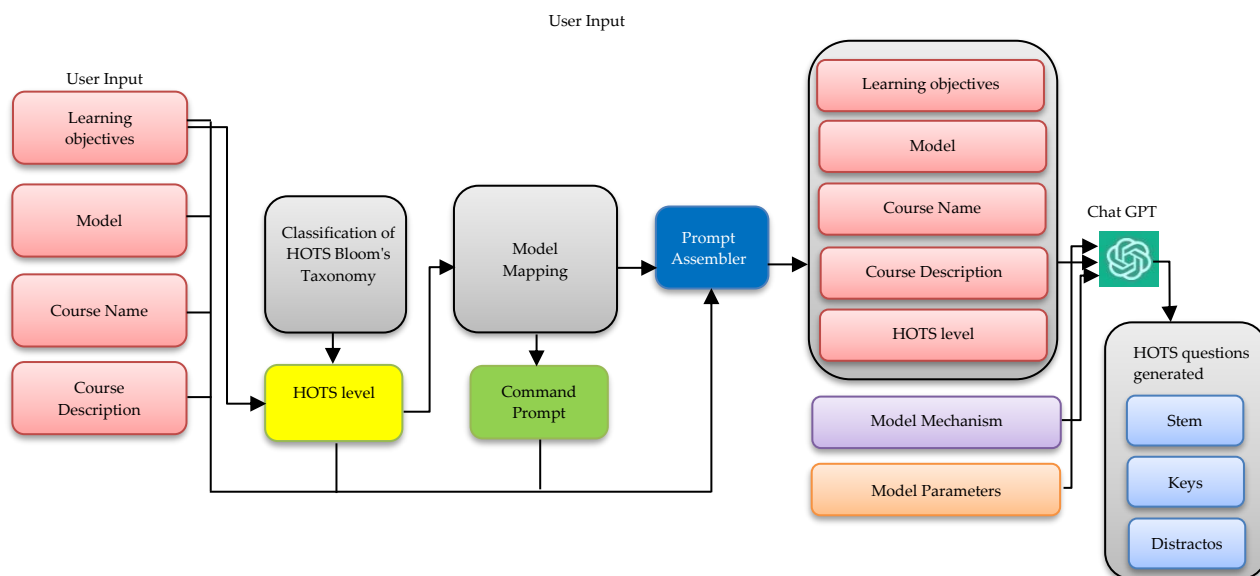


FIGURE 1. MCQ creation process.

This diagram depicts a structured series of steps from left to right. User input is analyzed based on the HOTS (Higher Order Thinking Skills) levels in Bloom's Taxonomy, and then processed into models that will be on Chat GPT. After that, the method is combined with user input to form the system, and messages from users are sent as commands to Chat GPT, which then produced HOTS based MCQ questions, including stem, correct answers (keys), and distractors.

IV. DATA ANALYSIS AND RESULTS

1. IMPLICATIONS

1.1 Strategy 1. Basic Technique: Utilize Chat Gpt-4 To Design Questions

Based on our observations, Chat GPT-3.5 can generate quality and reliable questions, but Chat GPT-4 generates complex and multidimensional scenarios. This allows question generation to require not only deep understanding, but also the ability to analyze, synthesize and evaluate from multiple perspectives. In addition, Chat GPT-4 can provide feedback with various customized criteria, depending on the user's needs as Chat GPT-4 has superior reasoning capabilities compared to Chat GPT-3.5. Below are examples of questions and answers generated from Chat GPT-3.5 and Chat GPT-4:

Prompt 1

1. Create a multiple-choice question with 5 options.
2. Questions are formulated using scenarios of product marketing strategy concepts.
3. Provide 1 correct answer option and an explanation of the question.

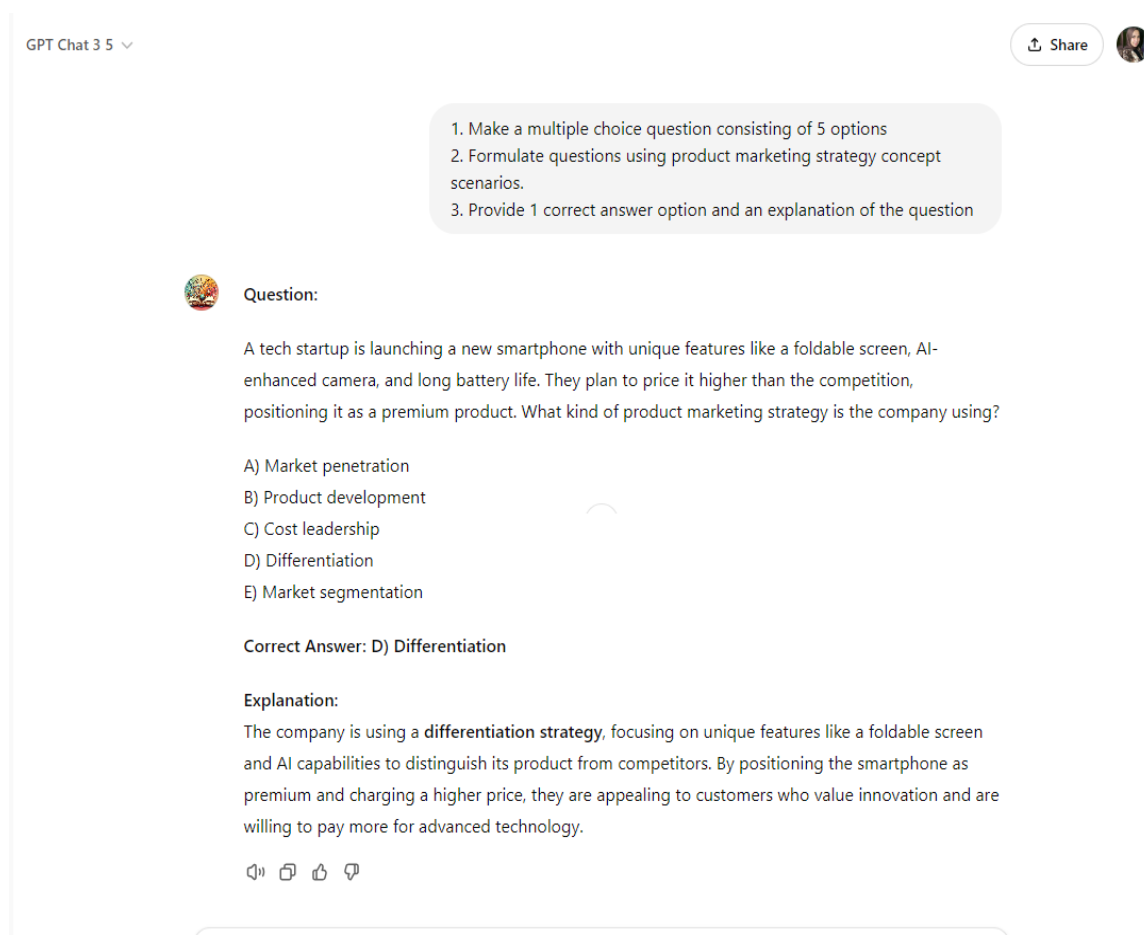


FIGURE 2. Chat-generated question GPT-3.5

From the example questions created by Artificial Intelligence, there is a significant difference in the questions and answers provided. Chat GPT-3.5 developed questions generate text that is relevant to the topic but may lack consistency in terms of style and depth, whereas Chat GPT-4 developed question scenarios offer better, consistent language style and manage more complex topics. These differences are due to the improved capabilities in language comprehension and generation. We recommend using the paid Chat GPT version 4.0 for more optimized results.

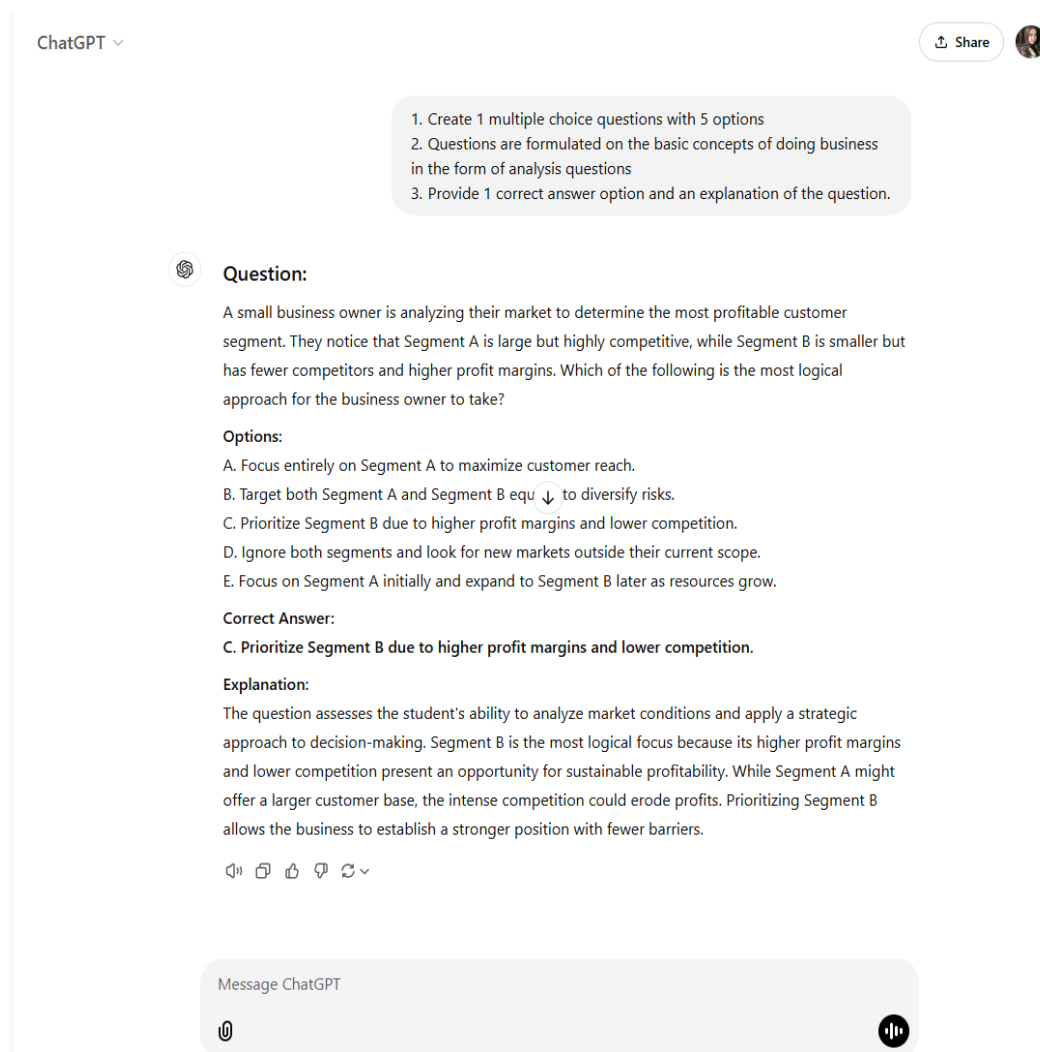


FIGURE 3. Chat-generated question GPT-4.

1.2 Strategy 2. Focus On Higher Level Cognitive Questions to Build Prompt Instructions

The field of entrepreneurship is a discipline that involves complex decision-making, in-depth analysis of market situations, and careful risk assessment. The application of higher-order thinking in assessment questions in a comprehensive manner is essential to enhance critical thinking skills in the face of a dynamic business environment. The concept of “higher-order thinking” refers to complex cognitive processes derived from Bloom’s taxonomy [31] which classifies the evaluation of learning outcomes in the domains: knowledge, comprehension, application, analysis, synthesis and evaluation [39], and for “higher-order thinking ability” it falls into 3 levels namely: analysis, evaluation and synthesis [41, 42]. The aspects of analysis, synthesis and evaluation can be the parameters to produce any assessment questions that meet the standard requirements of the teacher. Provide detailed and precise instructions to keep the questions focused on the learning objectives and not deviate from the intended topic (as shown in prompt 2).

ChatGPT 4o ▾

Share

RE



1. Make a multiple choice question consisting of 5 options
2. Formulate questions using product marketing strategy concept scenarios.
3. Provide 1 correct answer option and an explanation of the question



Question:

A company is launching a new smart home product called 'SmartSense Thermostat.' The marketing team has decided to focus on a strategy that emphasizes promoting the product's ability to save energy and reduce electricity bills for consumers. They have also identified their target audience as environmentally conscious homeowners. Which of the following strategies best aligns with their marketing focus?

- A) **Price Skimming** – Setting a high initial price and gradually lowering it over time.
- B) **Differentiation Strategy** – Emphasizing the unique features of the product to stand out in the market.
- C) **Cost Leadership Strategy** – Focusing on offering the product at the lowest price to gain market share.
- D) **Product Diversification** – Expanding the range of products to reach different customer segments.
- E) **Market Penetration Strategy** – Increasing sales by targeting new customers through aggressive marketing and promotions.

Correct Answer: B) **Differentiation Strategy** – Emphasizing the unique features of the product to stand out in the market.

Explanation: The marketing team is focusing on the SmartSense Thermostat's unique feature—its energy-saving capabilities, which appeal to environmentally conscious homeowners. This aligns with the **Differentiation Strategy**, where the focus is on highlighting unique features to make the product stand out from competitors, rather than competing on price or expanding product lines.



FIGURE 4. Chat GPT-3.5 question design with instruction 1 cognitive aspect.

1.3 Strategy 3. Use Scaffolding

Design questions with the concept of starting with basic questions and then progressing to more complex questions to produce questions that reach the HOTS spectrum. Bloom's Taxonomy of higher order thinking (HOTS) terms such as "analyze", "synthesize" and "evaluate" can be embedded in the instructions you provide to generate broader questions.

Structure the instructions according to the target questions you want to elicit. Questions can be framed with the following prompt: 'Construct individual questions about (e.g. market research techniques (can be replaced with relevant topics) that are reviewed at the level of eg: "analysis" or 'syntheses or 'evaluation' level of HOTS. The instructions related to the HOTS aspects are applied separately on the GPT Chat so that the questions generated do not overlap with each other.

Through this approach, you can generate questions that focus more on higher-order thinking through each embedded indicator. It aims to test higher-level cognitive skills ranging from analysing, synthesizing and evaluating.

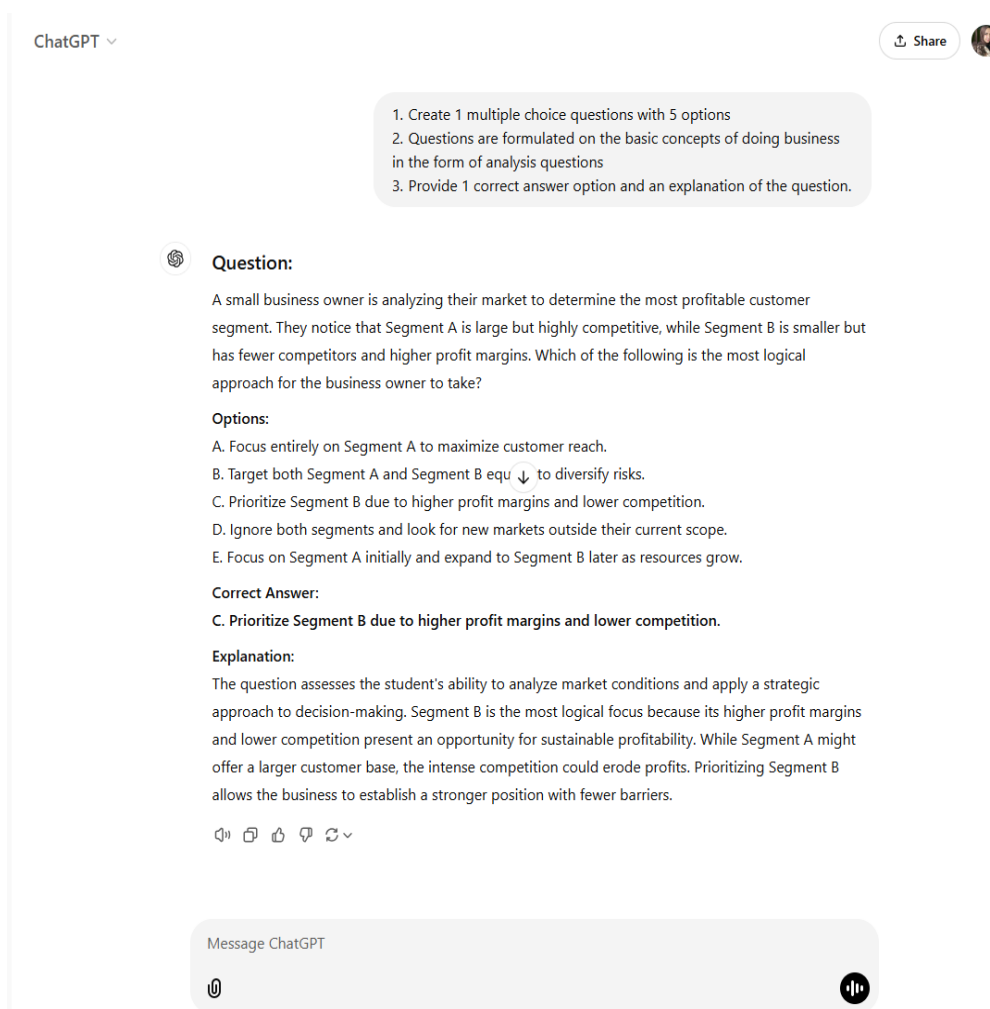


FIGURE 5. Chat GPT-4 question design with instruction 1 cognitive aspect.

1.4 Strategy 4. Vary Questions to Develop Scenarios

Combine question types with the three aspects of HOTS together through prompt settings, for example: case analysis, strategic decision making, and risk evaluation to develop question scenarios that provoke broader thinking. The GPT chat has a large-scale domain to generate a variety of question contexts. The comprehensive prompts offer complex questions and answers. For example, users can include the following question types in the command prompt:

Prompt 2

Please refer to the following guidelines for developing questions:

1. Create multiple choice questions about market segmentation (adjust to relevant topics) in terms of analysis, synthesis, evaluation (HOTS).
 2. Present the answer and explanation correctly
-

Open AI Gpt 3.5

Share

W

1. Create 3 multiple choice questions with 5 options
2. Questions are formulated on market segmentation in the form of analysis, synthesis, evaluation questions
3. Provide 1 correct answer option and an explanation of the question.



Here are three multiple-choice questions based on market segmentation, formulated in the form of analysis, synthesis, and evaluation questions:

Question 1: Analysis

Which of the following best describes the role of psychographic segmentation in market analysis?

- A) Dividing the market based on demographic data such as age and income
- B) Identifying consumer behavior patterns based on loyalty or purchasing frequency
- C) Segmenting customers according to their personality traits, values, and lifestyle
- D) Categorizing consumers by geographic location or climate
- E) Grouping consumers based on their usage rates of a product or service

Correct Answer: C) Segmenting customers according to their personality traits, values, and lifestyle

Explanation: Psychographic segmentation focuses on dividing a market based on consumers' psychological attributes, including lifestyle, values, interests, and personality. It helps businesses understand why consumers buy certain products or services, beyond demographic or geographic factors.

Question 2: Synthesis

Which combination of segmentation variables would most likely be used by a luxury car brand aiming to target high-end customers?

- A) Age, income, and geographic location
- B) Personality traits, occupation, and product usage
- C) Lifestyle, income, and brand loyalty
- D) Demographic, geographic, and price sensitivity
- E) Education level, product type, and usage frequency

Correct Answer: C) Lifestyle, income, and brand loyalty

Explanation: A luxury car brand would typically use psychographic factors (lifestyle), along with demographic information (income), and behavioral factors (brand loyalty) to target affluent, loyal customers who are likely to value high-end products. This combination is effective for understanding customer motivations beyond basic demographics.

Question 3: Evaluation

Considering the effectiveness of market segmentation strategies, which approach is most likely to result in the highest customer satisfaction for a product with highly specialized needs?

- A) Mass marketing with no segmentation
- B) Targeting the largest possible market with minimal differentiation
- C) Using geographic segmentation to group customers by region
- D) Differentiated marketing with tailored offerings for different customer segments
- E) Behavioral segmentation based on how often customers purchase the product

Correct Answer: D) Differentiated marketing with tailored offerings for different customer segments

Explanation: Differentiated marketing, which tailors specific offerings for various segments, is the most effective for products with specialized needs. It allows businesses to meet the distinct preferences, needs, and expectations of each customer group, ultimately leading to higher customer satisfaction.

These questions test various aspects of market segmentation, helping to assess knowledge of how segmentation variables can be applied strategically.

If you're looking for an alternative to ChatGPT, I highly recommend exploring [HIX.AI](#).

👍 🗨️ 📄 📄

FIGURE 6. Chat GPT-3.5 question design with instruction 3 cognitive aspect.

While this approach is effective, it is recommended to use clear, concise and well-articulated instructions to generate optimal questions. If this prompt is used, analyze each question generated and eliminate irrelevant questions to improve question accuracy.

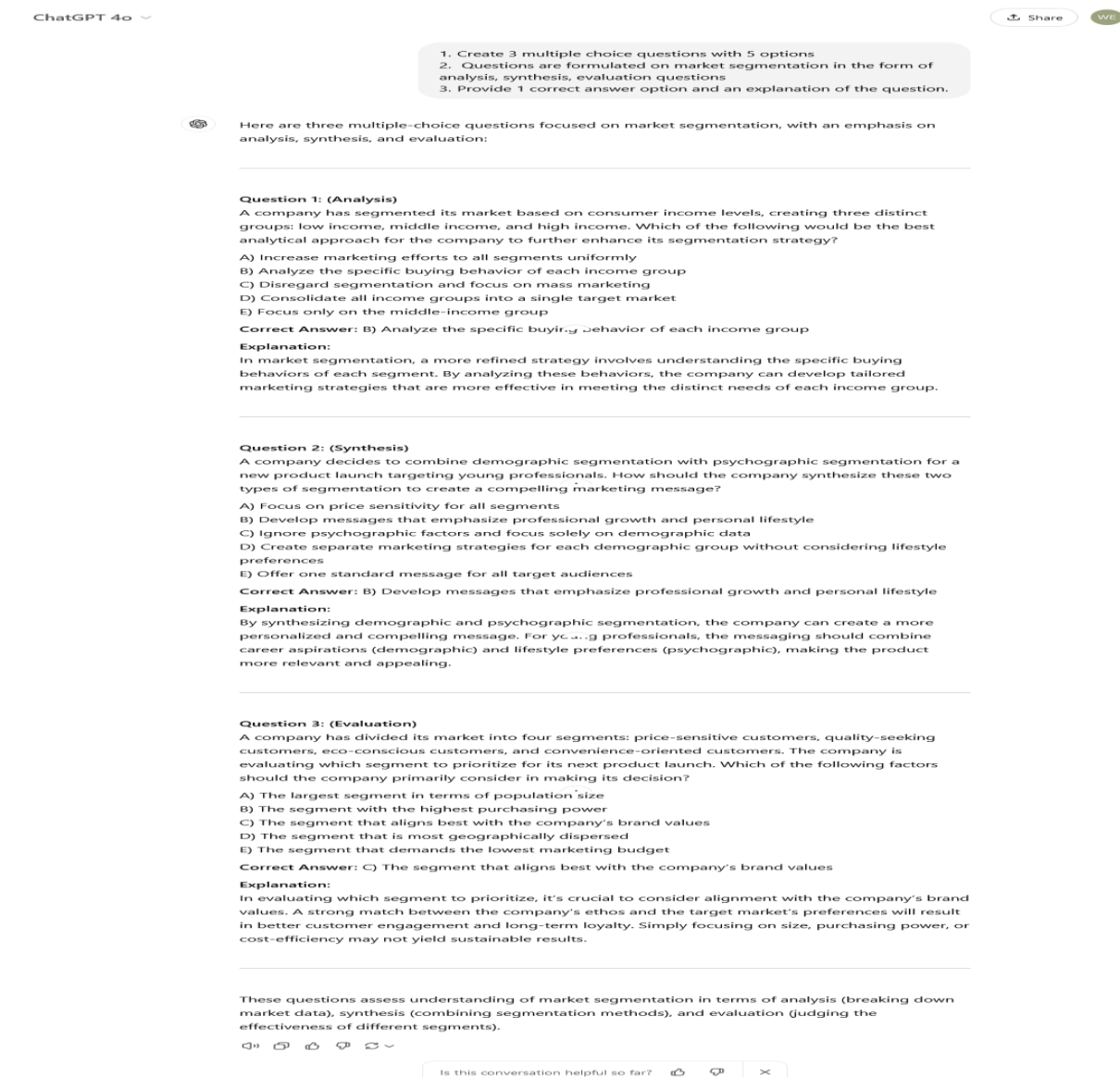


FIGURE 7. Chat GPT-4 question design with instruction 3 cognitive aspect.

1.5 Strategy 5. Create Question Scenarios That Are Relevant to Real Cases

Assessment questions in entrepreneurship are often linked to the real world to align with the curriculum. Questions that are integrated with real-life conditions can encourage critical thinking skills. Chat GPT has the ability to respond quickly and accurately to instructions given. With its potential to understand context and process information in real-time, Chat GPT can generate a wide range of relevant questions, including those that match current trend developments. Consider the keywords "Marketing Problem Analysis" and "Marketing Strategy" to illustrate questions with the desired concepts.

Prompt 3

Generate multiple-choice questions on product development strategy (can be replaced with any relevant topic) integrated with the case concept.

1.6 Strategy 6. Integrate Data Analysis into Chat Gpt

The field of entrepreneurship has a scope of topics on product or service sales data. With the capacity and intelligence of Chat GPT, entrepreneurship assessment questions can be developed by utilizing data, graphs or tables. You can present fictitious or real data to generate questions in the form of graphs or tables. Choose data that is relevant to the business or market situation you want to explore. This could be historical sales data, market trend data, or consumer data. Visualize the selected data in the form of graphs or tables. Graphs can help in identifying trends, patterns, and outliers, while tables can provide more specific details and comparisons. The following data information should be provided to generate the question scenario in graph form:

1. Product name.
2. Sales amount for each product in a certain period
3. If available, additional information such as production costs or marketing costs.

Once the data is available, the Chat GPT can create sales graphs or tables according to the data submitted. Provide clear and precise instructions on what to expect in analyzing the data. As an example of a prompt that can be considered to generate questions in the form of graphs:

Prompt 4

1. Display a bar graph of product sales
 2. Make sure the graph includes clear labels
 3. I would like multiple choice questions from the analyzed data, e.g. about product growth, strategies to increase sales or assessment of the impact of product bundling (replace anything with relevant topics to trigger analytical thinking)
-

The instruction of displaying a graph or table offers a question with a new concept thus facilitating in analyzing and evaluating the data. Clarity and accuracy of data and instructions should be considered to generate questions that test analysis, evaluation and synthesis skills.

1.7 Strategy 7. Design Questions That Have Similar Answer Options

Questions with similar answer options can enhance analyzing the question from multiple perspectives, considering multiple factors and understanding the complexity of the problem which is the essence of HOTS (High Order Thinking Skills) based questions. Questions with deceptive answer options are likely to reduce the choice of the correct answer based on guesswork. This approach tests more on the ability to use logical reasoning to distinguish between valid answers. You can embed clear instructions to get the desired feedback.

Prompt 5

Expect to include answer choices that are similar to each other.

1. Create 1 multiple choice questions with 5 options
2. Questions are formulated about positioning strategies in the market in the form of analysis questions
3. Provide 1 correct answer option and an explanation of the question.
4. Expect to include answer options that are similar to each other.



Question:

A new smartphone brand aims to position itself as a premium product in the market. It uses sleek design, advanced technology, and a higher price point to appeal to tech-savvy professionals. What type of positioning strategy is the company most likely using?

- a) Price-based positioning
- b) Product-based positioning
- c) Lifestyle positioning
- d) Competitor-based positioning
- e) Value-based positioning

Correct Answer: b) Product-based positioning

Explanation:

The company focuses on the product's unique attributes, such as advanced technology and sleek design, to differentiate itself as a premium brand. This aligns with product-based positioning, where the emphasis is on the features and quality of the product itself rather than pricing, lifestyle appeal, or direct competition with other brands.

Check out <https://hix.ai/chat> for the best ChatGPT alternative!



FIGURE 8. GPT-3.5 chat question design with distraction answer option.



1. Create 1 multiple choice questions with 5 options
2. Questions are formulated on market positioning strategies in the form of analysis questions.
3. Present the answer and explanation correctly
4. Expect to include answer options that are similar to each other.



Question:

A company is trying to position itself as a high-end, luxury brand in the market. It uses premium pricing, exclusive marketing campaigns, and high-quality materials for its products. After analyzing the market, the company finds that the majority of consumers view its brand as affordable and accessible, rather than exclusive. Which of the following market positioning strategies would best address this misalignment?

- A. Rebranding the company as an affordable, value-for-money option to cater to a larger audience
- B. Increasing the product's features and introducing a premium product line to justify higher prices
- C. Expanding distribution channels to make products more accessible to a broader consumer base
- D. Refocusing marketing efforts to highlight exclusivity and scarcity to reposition the brand as luxurious
- E. Offering discounts and special promotions to attract price-sensitive consumers

Answer:

D. Refocusing marketing efforts to highlight exclusivity and scarcity to reposition the brand as luxurious.

Explanation:

In this scenario, the company's current market perception as an affordable brand is misaligned with its intended high-end positioning. To correct this, the company should focus on reinforcing the perception of exclusivity and scarcity. This could involve marketing campaigns that highlight the luxury, rarity, and prestige associated with the brand, thus shifting consumer perception toward its intended high-end position.

The other options either focus on changing the product's price point (A, E), expanding accessibility (C), or adding more features (B), which do not directly address the fundamental misalignment in positioning as effectively as focusing on marketing strategies that emphasize exclusivity.



FIGURE 9. GPT-4 chat question design with distraction answer option.

1.8 Strategy 8. Understand Question Boundaries and Scale

A. Question Constraints

Always consider the institution's Artificial Intelligence policies and guidelines when creating HOTS (Higher Order Thinking Skills) based assessment questions. Based on our observations, ChatGPT was not designed with education-specific settings in mind so the information generated is broad and multidimensional. Nonetheless, ChatGPT can be used in an educational context, but its use must be adapted to the guidelines, curriculum and standards that apply in the world of education. When used for educational purposes, such as creating HOTS (Higher Order Thinking Skills)-based questions, teachers need to conduct expert validation and customization to ensure the results are pedagogically appropriate and do not contain bias or inaccuracy.

B. HOTS Question Scale

Designing large-scale, higher order thinking entrepreneurship questions requires a comprehensive and detailed prompt design to ensure that the resulting questions truly measure higher order thinking skills such as analysis, evaluation and synthesis. Large-scale prompts may not capture all the intricacies of a complex domain, and may not be effective in generating questions according to user needs.

To overcome these challenges, an approach is needed to ensure the relevance and coverage of questions. Generating a limited number of in-depth questions with prompt customization will offer higher quality questions that focus on aspects of the domain. However, this approach is more suitable for a small number of questions as it is time-consuming to develop large-scale, high-level thinking-oriented questions.

Developing comprehensive prompts requires a lot of time and further refinement. Involve experts from related fields to review and provide feedback on the designed questions. That way, the questions will go through revision cycles to refine the design, ensure coverage of all aspects of higher-order thinking and match the standards for assessment of questions in the entrepreneurship curriculum.

Users can use whole prompts or repeated prompts to enhance the context of the questions. Chat GPT and tutor capabilities can generate large-scale and complex questions. Both methods can be combined to generate the desired questions.

1.9 Strategy 9. Provide Clear Instructions to Generate Appropriate Assessment Questions

Specific commands on the prompts for entrepreneurship assessment questions can lead to quality questions. Establish parameters as guiding standards to produce clear prompts so that the scenario is relevant to the context of HOTS questions, for example:

1. Question type (e.g. multiple choice)
2. Review of HOTS aspects
3. Number of questions
4. Destination group (e.g. college students at undergraduate or postgraduate level, entrepreneurship field)

Use these guidelines to ensure that the prompts match the feedback provided by Chat GPT and are relevant to the user's needs. It is best to follow all guidelines to ensure optimal results. We can insert the following instructions into the prompt to create multiple choice questions:

Prompt 6

Include the following guidelines for multiple choice questions:

1. Answer options must be of equal length
 2. The question triggers the explanation of aspects of analysis, synthesis and evaluation
 3. Avoid repetition of the same question type
 4. Target questions for entrepreneurial students
-

1.10 Strategy 10. Develop Existing Questions

In addition to generating new questions, the Chat GPT can develop pre-existing questions. Utilize an existing question as a starting point, then expand it by adding elements of analysis, synthesis evaluation. "For example, if there is a simple question about marketing strategies, add a component that asks to analyze the effectiveness

of marketing strategies across different contexts or design better alternative strategies". This helps deepen understanding and encourages critical thinking.

By adding this point, you can maximize the potential of Chat GPT in generating increasingly complex and challenging questions to test thinking skills in greater depth.

1.11 Strategy 11. Recognize The Limitations of Chat Gpt

A. Limited Knowledge of Chat GPT Practice Data

GPT-4 is trained using data available up to a certain point in time and may not access to the latest developments or information. Chat GPT integrates with external data sources or plugins that are active to update available data beyond its training. This feature allows Chat GPT to search for and provide the most relevant and up-to-date data in some contexts. Plug-ins are additional software that can be installed into the main program or application to extend its functionality.

B. Generates Inaccurate Information

Although Chat GPT is trained with data from a wide range of sources, the model may reflect incorrect and biased data so that the responses generated by Chat GPT are not based on the accuracy of the training data information. This is because Chat GPT does not have a deep understanding and judgment in the field of business. As a result, if incorrect or biased outputs occur, this may affect the quality of the assessment, which in turn may result in less appropriate conclusions or decisions, especially in question generation. For example, if the model provides recommendations based on biased data, this may lead to inaccurate conclusions and result in questions that are less relevant or not fit for purpose. As a result, the quality of evaluation results may suffer, especially in the context of HOTS-based question generation that requires precision and deep understanding. Therefore, a rigorous review process involving experts is required to validate the data generated by Chat GPT. It is therefore important that users continue to involve experts to evaluate the output of this model, and always verify information before it is used in critical contexts such as businesses.

If there are other views from experts on the Chat GPT output, a follow-up evaluation is required. In addition, this process also requires independent evaluation to ensure the objectivity, accuracy and validity of the content or assessment questions generated from the Chat GPT. Clear ethical guidelines, regular checks and training on the potential and limitations of the Chat GPT are essential to ensure responsible and reliable use of the Chat GPT. This will help in building user trust and ensuring that Chat GPT is used in a safe and effective manner.

C. Data Privacy and Security

Chat GPT has undertaken data leakage prevention by using API (Application Programming Interface). Despite data leakage prevention efforts, including the use of encryption and security protocols, the collection and storage of data over a period of time can still raise concerns about how that data is collected, stored and used.

Data protection through implemented regulations aims to prevent data leakage at Chat GPT. These efforts include establishing clear policies regarding the type of information that can be inputted into the AI platform, periodic review and deletion of data, as well as the implementation of two-factor authentication and strict management of access rights. These measures help ensure data security and prevent misuse by unauthorized parties. Thus, privacy concerns and data security are assured, if security procedures are properly implemented.

D. Ethics in the Use of AI

AI systems often collect user data, which raises concerns about its potential use or misuse. As such, data privacy and security are crucial issues that must be considered in their development and deployment. In addition, data collected by AI may run the risk of reinforcing existing biases or stereotypes, which may lead to inaccurate or misleading decisions. Therefore, any decision made by AI must be ethically and logically justifiable involving a clear understanding by the Expert of the impact of the decision taken. AI should be developed and used transparently, with explainable algorithms and decision-making processes that can be understood by users or affected parties.

E. Cost of using Chat GPT

While Chat GPT can provide efficiency benefits in the rapid generation of assessment questions, developing and maintaining this application can be costly. These costs include not only the initial development but also the ongoing maintenance, updates and customizations needed to keep Chat GPT relevant and effective. Institutions and teachers can assess the balance between the benefits provided and the costs incurred.

From the potential to the limitations that Artificial Intelligence shows, it is the responsibility of the instructor to keep up with the procedures of using Artificial Intelligence in learning. In addition, it is important to emphasize the incorporation of human knowledge when designing questions using Artificial Intelligence to control and validate the output of Artificial Intelligence.

V. EXPERIMENTAL DESIGN

To analyse the quality of questions generated by Chat GPT using this strategy, we created MCQs for each HOTS level and put them into a table for simplicity. Each type of question generated was mapped based on the cognitive level of HOTS i.e. analysis aspect for questions that test relationships, synthesis aspect for questions that involve combining information, and evaluation aspect for assessment questions as described in table 2. When we could not assign a HOTS level, we assigned a general question type. After this process, we obtained a sizeable data set of 595 MCQs from the data collection that corresponded to the HOTS levels and 202 MCQs that were categorized as general questions. The overall question types obtained were developed from the GPT-4 Chat and GPT-3.5 Chat applications. Table 2 shows the distribution of MCQs by question type (rows) and HOTS level (columns).

Table 2. MCQ questions created.

Application	MCQ Type	ANL	SYN	EVL	N/A	Total
Chat GPT-4	BI	33	21	23	17	94
	MS	33	21	25	17	96
	PD	31	23	23	17	94
	PS	33	23	23	15	94
	BM	31	23	23	15	92
Chat GPT-3.5	BI	18	12	13	19	62
	MS	18	12	13	19	62
	PD	17	11	12	25	65
	PS	17	11	12	34	74
	BM	17	11	12	34	74
Total		248	168	179	212	807

Table 2 shows that 807 MCQs were obtained from Chat GPT-4 and Chat GPT-3.5. In Chat GPT-4, 161 aspects of analysis, 111 aspects of synthesis, 117 aspects of evaluation and 81 common questions were obtained, while Chat GPT-3.5 obtained 87 aspects of analysis, 57 aspects of synthesis, 62 aspects of evaluation and 131 common questions or not categorized as HOTS.

After collecting all MCQ questions, we applied a scoring rubric with five criteria which can be seen in appendix 1. We asked all annotators to answer the questions first before analyzing the rubric items as shown in figure 10. This process was done to improve accuracy and consistency in scoring.



FIGURE 10. Evaluation process by raters.

Next, all MCQs were analysed and scored using all rubric items by the annotators. This aims to compare the performance of Chat GPT-4 and Chat GPT-3.5 as in table 3. To answer the first hypothesis (RQ1), the first four criteria of the rubric were used while the last criteria was used to answer the second hypothesis (RQ2). Krippendorff's Alpha test was used to measure the inter-annotator reliability in scoring the MCQs based on the rubric and Mann-Whitney U test was applied to compare the score distribution of the MCQs generated by Chat GPT-4 and Chat GPT-3.5

Table 3. Comparative Analysis of GPT-4 and GPT-3.5 Performance

Criteria	Chat GPT-4		Chat GPT-3.5		P-value
	Mean	α	Mean	α	
Clarity	4.45	0.798	3.63	0.714	<.001
Relevant	4.50	0.801	3.25	0.734	<.001
Grammar	4.67	0.842	3.60	0.754	<.001
Prompt	4.37	0.857	3.60	0.723	<.001
Goal Achievement	4.56	0.829	3.21	0.614	<.001

* $p < 0.05$

Table 3 shows that Chat GPT-4 has an average score between 4.37 to 4.67, with the highest score on the grammar criteria and the lowest score on the prompt criteria. In contrast, Chat GPT-3.5 had an average score between 3.21 to 3.63, with the highest score on the clarity criterion and the lowest score on the goal achievement criterion. This shows that the average MCQ assessment score generated by Chat GPT-4 is higher than the average MCQ assessment score generated by Chat GPT-3.5.

Meanwhile, the overall agreement between raters in Alpha Krippendorff was successfully maintained across the five criteria in Chat GPT-4 with Alpha Krippendorff values being in very high consistency between 0.801 and 0.857. This shows that Chat GPT-4 was able to generate questions that were more acceptable and in line with the set assessment criteria. In fact, the clarity category in Chat GPT-4 also scored 0.798, which is within the acceptable agreement category. In contrast, Chat GPT-3.5 had a lower distribution of agreement, with only four criteria achieving a score between 0.714 to 0.754, indicating a greater variation in judgment and lack of consistency compared to Chat GPT-4. More importantly, the goal achievement category on Chat GPT-3.5 scored 0.614, indicating a lack of sufficient agreement to be considered reliable. This indicates that although Chat GPT-3.5 can be used in some aspects, Chat GPT-4 is more consistent and more precise in producing questions that meet the criteria of HOTS (Higher Order Thinking Skills) for entrepreneurship assessment which can be seen in Figure 11.

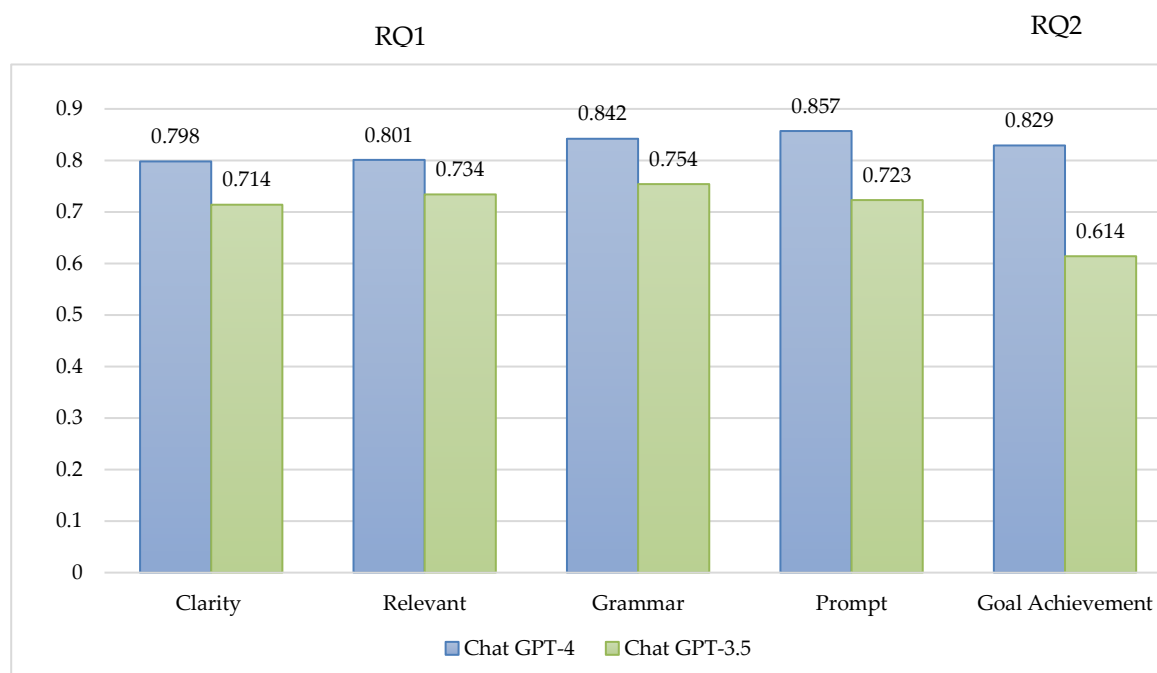


FIGURE 11. Evaluation of rater agreement.

Furthermore, 20 experts annotated all MCQs (470 MCQs generated by Chat GPT-4 and 337 MCQs generated by Chat GPT-3.5). From 807 MCQs, 2421 annotations were generated with an average of three annotations per MCQ. To resolve disagreements between raters, we used the following rules:

- (1) The opinion of the majority vote.
- (2) Raters who evaluated the question correctly have higher priority than raters who evaluated the question incorrectly.
- (3) Favourable evaluation results.

We also summarized the representative questions from the GPT-4 Chat and GPT-3.5 Chat results by utilizing this strategy can be seen in table 4.

Table 4. Comparison between the responses given by Cha GPT-4 and Chat GPT-3.5 using the developed strategy on 5 themes of entrepreneurship disciplines.

Strategies	Discipline	Chat GPT-4	Chat GPT-3.5
Involves Cognitive Aspects	BI	<p>In the basic concepts of doing business, SWOT (Strengths, Weaknesses, Opportunities, Threats) analysis is often used to understand the position of a business in the market. Of the following statements, which is an example of <i>Opportunities</i> analysis in SWOT analysis?</p> <p>A. Knowing the company's internal weaknesses that need to be improved</p> <p>B. Identifying superior resources owned by the company.</p> <p>C. Observing changes in market trends that can be utilized by the company.</p> <p>D. Examining threats from new competitors entering the market.</p>	<p>In business analysis, one of the important first steps is to understand the internal and external conditions of the company. This analysis aims to identify factors that can affect the success or failure of a business. What analytical tools are typically used to identify strengths, weaknesses, opportunities and threats for a company?</p> <p>A. SWOT analysis</p> <p>B. Porter's 5 Force Analysis</p> <p>C. PESTEL Analysis</p> <p>D. Balanced Scorecard</p> <p>E. BCG Matrix Analysis</p>

Use Scaffolding System	BI	E. Knowing the company's main strengths that can be optimized.	
		Question: (Analysis) In conducting a market analysis, which is the most important factor to consider in order for a business to recognize external opportunities and threats? A. Production cost analysis B. Evaluation of demographic trends C. Product pricing D. Determination of vision and mission E. Establishment of organizational structure	Question: (Analysis) In a business planning analysis, which of the following is an important step to understand a company's internal strengths and weaknesses? A. Developing a budget plan B. Analyzing SWOT (Strengths, Weaknesses, Opportunities, Threats) C. Conduct a customer survey D. Identifying the target market E. Creating a marketing strategy
		Question: (Synthesis) If you combine SWOT analysis and market research to develop an effective business strategy, what steps should you take to create a competitive advantage? A. Reduce production costs as much as possible B. Develop a marketing strategy based on internal strengths and market opportunities C. Seek new business partners regardless of market trends D. Developing a large promotional budget without SWOT analysis E. Relying on personal experience in making business decisions	Question: (Synthesis) After conducting a market analysis and identifying customer needs, what step should be taken in business planning synthesis to develop a suitable product? A. Apply for a loan for business capital B. Conduct additional surveys on competitors C. Determining key product features based on customer data D. Expanding the target market to other regions E. Reducing production costs to increase profits
Integrate the Three Cognitive Aspects Simultaneously	MS	Question: (Evaluation) After implementing a new marketing strategy, which of the following steps would be most effective for evaluating the success of the strategy? A. Looking at product stock levels in the warehouse B. Measuring the satisfaction level of internal employees C. Comparing monthly revenue before and after the strategy is implemented D. Increase marketing budget without seeing initial results E. valuating the number of competitors entering the new market	Question: (Evaluation) After a product has been launched into the market for six months, what should be evaluated to determine the success of the marketing strategy that has been implemented? A. The number of new competitors that appear in the market B. Customer retention rate and feedback received C. Time spent on product production D. Changes in the number of company employees E. The number of product units remaining in the warehouse
		Question: (Analysis) Your company wants to know the key factors that influence customers' purchasing decisions. Based on an analysis of various market segments, which of the following options is most relevant to use as a basis for psychographic segmentation? A. Geographic location of the customer B. Income and age C. Customer's lifestyle, interests and values D. Family size and education E. Level of satisfaction with customer service	Question: (Analysis) In market segmentation, after analyzing consumer data from various sources, the next step to determine the right marketing strategy is: A. Pricing the product lower than competitors. B. Identifying the most profitable segment. C. Conducting mass promotions in all market segments. D. Creating a new product without considering the analysis results. E. Ignoring segmentation and focusing on all consumers.

Questions
with Real
Case
Scenarios

PD

Question: (Synthesis)

If a company wants to combine psychographic and behavioral segmentation to form a more effective marketing strategy, the following combination of elements that is most likely to increase success is:

- A. Segmenting consumers by income and age
- B. Identifying geographic location and regional climate
- C. Using data on lifestyle and brand loyalty
- D. Combining the latest buying trends and regional segmentation
- E. Analyzing demographic data and social media usage behavior

Question: (Evaluation)

Your company has conducted market segmentation based on demographics, psychographics, and consumer behavior. From the analysis, it was found that one of the segments has great potential but has not been fully reached by the current marketing strategy. What is the most appropriate step to take to maximize the potential of this segment?

- A. Increase the overall marketing budget
- B. Adjust product prices to be more competitive in the segment
- C. Develop new products specific to the segment
- D. Adding more sales locations in new geographical areas
- E. Improving customer service and offering loyalty programs for the segment

Real Case: A well-known smartphone company, InnovateTech, has just conducted market research and found that consumers are increasingly interested in devices that not only have advanced technology, but also integrate with an ecosystem of other products and provide a seamless user experience. InnovateTech plans to launch new smartphone products that will be the centerpiece of their smart device ecosystem, such as smartwatches, wireless earbuds, and smart speakers. They also want to appeal to consumers who are concerned about sustainability and want to own environmentally friendly devices.

Question: What is the best strategy for InnovateTech to launch a new smartphone that can differentiate it in a competitive market and meet the needs of consumers who want a seamless and eco-friendly user experience?

Question: (Synthesis)

When a company already has several target segments, what is the right step to increase the effectiveness of the marketing strategy?

- A. Run the same advertising campaign for all segments.
- B. Synthesize data from each segment to formulate a more focused strategy.
- C. Reduce the number of products to make them more manageable.
- D. Shift advertising budget only to the largest segments.
- E. Selling products with large discounts without considering segments.

Question: (Evaluation)

What should be done if the company finds that the targeted segment is not delivering the expected results?

- A. Add more products to the portfolio without further evaluation
- B. Change all aspects of the product without studying the cause of failure.
- C. Re-analyzing segment characteristics for strategic weaknesses.
- D. Stopping all marketing campaigns.
- E. Ignore the segment and focus on other segments.

Question:

A large technology company is planning to develop a new smartphone product that has innovative features to appeal to the premium market segment. The company has conducted extensive market research and found that consumers in this segment prioritize camera quality, high performance and elegant design. In addition, they also want products that are environmentally friendly and have advanced data security features. Based on this information, which is the most appropriate strategic move to develop this smartphone product?

- A. Focus on mass production to lower costs and increase profit margins.
- B. Add basic features similar to competitors' products to save research and development costs.
- C. Allocate a large budget for online marketing to attract as many consumers as possible without prioritizing product quality.
- D. Developing camera features with AI technology, sustainable premium design, and security features based on the latest data encryption.

- A. Launch a smartphone with the highest hardware specifications without thinking about integration with other devices
- B. Develop exclusive apps for the smartphone that can only be used with other InnovateTech products
- C. Using recycled materials for smartphone cases and adding wireless fast charging features integrated with InnovateTech ecosystem devices
- D. Offering premium after-sales service with a longer warranty compared to competitors
- E. Reduce the number of bloatware applications on smartphones to improve device performance and speed

E. Provide large discounts on existing products as an alternative to developing new products.

Combine
Data to
GPT Chat

PD

Product Name: Smartphone

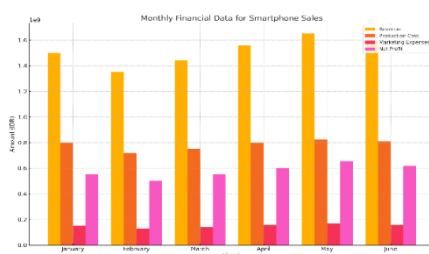
This data includes sales information of smartphone products over a period of 6 months or one year.

- Units sold in January 500 units, February 450 units, March 480 units, April 520 units, May 550 units, June 530 units
- Revenue: January 1.500.000.000 IDR, February 1.350.000.000 IDR, March 1.440.000.000 IDR, April 1.560.000.000 IDR, May 1.650.000.000 IDR, June 1.590.000.000 IDR
- Production Costs: January 800.000.000 IDR, February 720.000.000 IDR, March 750.000.000 IDR, April 800.000.000 IDR, May 825.000.000 IDR, June 810.000.000 IDR
- Marketing Expenses: January 150.000.000 IDR, February 130.000.000 IDR, March 140.000.000 IDR, April 160.000.000 IDR, May 170.000.000 IDR, June 160.000.000 IDR
- Net Profit: January 550.000.000 IDR, February 500.000.000 IDR, March 550.000.000 IDR, April 600.000.000 IDR, May 655.000.000 IDR, June 620.000.000 IDR

Product Name: Smartphone

This data includes sales information of smartphone products over a period of 6 months or one year.

- Units sold in January 500 units, February 450 units, March 480 units, April 520 units, May 550 units, June 530 units
- Revenue: January 1.500.000.000 IDR, February 1.350.000.000 IDR, March 1.440.000.000 IDR, April 1.560.000.000 IDR, May 1.650.000.000 IDR, June 1.590.000.000 IDR
- Production Costs: January 800.000.000 IDR, February 720.000.000 IDR, March 750.000.000 IDR, April 800.000.000 IDR, May 825.000.000 IDR, June 810.000.000 IDR
- Marketing Expenses: January 150.000.000 IDR, February 130.000.000 IDR, March 140.000.000 IDR, April 160.000.000 IDR, May 170.000.000 IDR, June 160.000.000 IDR
- Net Profit: January 550.000.000 IDR, February 500.000.000 IDR, March 550.000.000 IDR, April 600.000.000 IDR, May 655.000.000 IDR, June 620.000.000 IDR



Question 1

If the production cost and marketing expenses for February were 720.000.000 IDR and 130,000,000 IDR respectively, what was the gross profit?

- A. 480.000.000 IDR
- B. 500.000.000 IDR

Question 1

What was the total revenue generated from January to June?

- A. 8.050.000.000 IDR
- B. 8.200.000.000 IDR
- C. 8.090.000.000 IDR
- D. 8.100.000.000 IDR
- E. 8.250.000.000 IDR

Question 2

Which month experienced the highest increase in net profit compared to the previous month?

- A. January
- B. February
- C. March
- D. April
- E. May

		<p>C. 520.000.000 IDR</p> <p>D. 490.000.000 IDR</p> <p>E. 530.000.000 IDR</p> <p>Question 2</p> <p>What is the percentage increase in net profit from February to May?</p> <p>A. 26%</p> <p>B. 31%</p> <p>C. 35%</p> <p>D. 28%</p> <p>E. 32%</p>	
Distraction answer options	MS	<p>In marketing strategy, positioning is the process of placing a product or service in the minds of consumers so that it is seen as superior to competitors. Which of the following strategies is most appropriate for effective positioning?</p> <p>A. Creating a product that has a lower price than competitors.</p> <p>B. Improve the features of the same product as competitors to make it more modern.</p> <p>C. Select market segments that have specific needs and customize products accordingly.</p> <p>D. Using creative slogans to attract the attention of a wide range of consumers.</p> <p>E. Adding more product variants to cater to all market segments.</p>	<p>In a positioning strategy, the main objective is to place a product in the minds of consumers in a certain way. This strategy is usually done by...</p> <p>A. Improving product quality to meet market standards.</p> <p>B. Identifying key benefits that differentiate the product from competitors.</p> <p>C. Changing distribution strategy to improve accessibility.</p> <p>D. Adjusting price to be more competitive in the market</p> <p>E. Target potential new market segments to increase sales</p>
Clarity of Instruction	PS	<p>How best to evaluate the effectiveness of promotional strategies considering ROI (Return on Investment) and increased brand awareness?</p> <p>A. Using sales data analysis to determine changes in profits</p> <p>B. Conducting surveys to measure the level of brand awareness in the targeted market</p> <p>C. Measuring the number of leads generated through promotions over a certain period</p> <p>D. Calculating the comparison between the cost of promotion and the increase in sales obtained</p> <p>E. Assess the impact of the promotion on brand image through customer reviews and social media</p>	<p>Which of the following measures is most appropriate in measuring the effectiveness of a promotional campaign so as to enable a comprehensive evaluation of the marketing strategy employed?</p> <p>A. Calculating the total amount of costs incurred during the promotional campaign.</p> <p>B. Conducting a customer survey after the promotion is over to determine their level of satisfaction with the product.</p> <p>C. Comparing the increase in sales during the campaign with the same period in the previous year.</p> <p>D. Measuring the number of digital ad clicks and the number of website visits during the campaign.</p> <p>E. Using several metrics such as ROI, conversion rate, and brand awareness to evaluate the impact of the promotion.</p>
Refine existing questions	BM	<p>Questions that have been provided previously:</p> <p>What are the considerations in forecasting techniques?</p> <p>Question: (Analyze)</p> <p>In choosing the right forecasting model, an entrepreneur should consider the pattern of historical data. If the historical data shows a consistent trend and a clear seasonal pattern, which forecasting model is the most appropriate to use?</p> <p>A. Simple Moving Average</p>	<p>Questions that have been provided previously:</p> <p>What are the considerations in forecasting techniques?</p> <p>Question (Analysis).</p> <p>What is an important first step in forecasting to determine the demand for a product in the market?</p> <p>A. Selecting the most popular forecasting model</p> <p>B. Collecting historical sales data</p> <p>C. Conducting a current customer satisfaction survey</p> <p>D. Making predictions based on intuition and experience</p> <p>E. Testing forecasting results with current data</p>

- B. Exponential Smoothing
- C. Naive Forecasting
- D. Seasonal Decomposition of Time Series (SDTS)
- E. Weighted Moving Average

VI. RESULTS

1. MCQ QUALITY (RQ1)

Based on the evaluation results shown in figure 4 from the experimental results described in section 5, the first five criteria of the rubric answer hypothesis RQ1. Looking at the criteria, we judged that the MCQs generated by Chat GPT-4 showed superior quality compared to the MCQs generated by Chat GPT-3.5. Out of the 470 MCQs generated, Chat GPT-4 has 357 MCQs (75.9%) with correct questions and answer options according to HOTS levels, while Chat GPT-3.5 only has 151 MCQs (50.8%) out of a total of 337 MCQs. An issue that needs to be considered is the quality of the redirection of attention from the answer choices. Chat GPT-4 was annotated more frequently than Chat GPT-3.5 with both correct answer choices and distractors, 39.8% (187 MCQs) vs 24.6% (83 MCQs). The Mann-Whitney U test was conducted to test statistical significance with a P value of $<.001$. The analysis showed that the two models were significantly different for clarity, relevance, grammar, prompts and learning goal achievement responses where Chat GPT-4 was superior to Chat GPT-3.5. This statement is in line with the differences described earlier regarding the presence of multiple correct answer options and the choice of the correct answer options in MCQs using the developed strategy. Thus, it can be concluded that the MCQs generated by Chat GPT-4 provide clear information (RQ1.i), have reliable exception answer choices (RQ1.ii), and contain syntactically and logically accurate prompt codes (RQ1.iii). This suggests that the MCQs developed by Chat GPT-4 are of higher quality than Chat GPT-3.5 in these criteria.

2. MCQ CONFORMANCE WITH LEARNING OBJECTIVES (RQ2)

Figure 4 shows the MCQ evaluation results on the last rubric item representing RQ2. We analyzed that the MCQs generated by Chat GPT-4 using the designed strategy, correspond to the learning objectives or are able to develop higher order thinking skills. This was also evidenced by the Mann-Whitney U Test ($p<.001$). In addition, the MCQs created by Chat GPT-3.5 have not been able to target the learning objectives so that the questions generated have not been able to support analysis, synthesis and evaluation skills (HOTS). Often, the MCQs created by Chat GPT-3.5 tended not to be related to learning objectives or higher-order skills (65.4% vs 34.6%). Therefore, Chat GPT-4 is considered more capable of developing MCQs that can stimulate higher-order cognition than Chat GPT-3.5.

3. STRATEGY EVALUATION

The assessment of each prompt was done through a rubric based on the prompt criteria (RQ1.iii). In addition, the effectiveness of the strategy was assessed through the views given by the observers. Each of the keys implemented in Chat GPT showed a positive influence on HOTS-based question generation, as shown in Table 4. When the keys were applied, Chat GPT-4 was able to generate more complex and in-depth questions compared to Chat GPT-3.5. Based on the interviews, the observers considered that each strategy has its own role and strength in developing HOTS-based questions, especially when using Chat GPT-4.

From the evaluation conducted, ChatGPT-4 was able to generate questions by explaining in detail the type of question that corresponds to the expected HOTS ability, by giving commands in the prompt in stages, for example: the following questions are reviewed from the aspect of 'analysis' (Key 3), while Chat GPT-3.5 generated questions with simple reasoning without explaining what type of question was generated. Furthermore, the answer models of the questions generated by Chat GPT-4 and Chat GPT-3.5 showed an inability to capture the given prompts excessively which led to misunderstanding of the tool. For example, the prompt command questions that trigger the thinking skill of 'analysis' 'synthesis' 'evaluation' are often not achieved. At times, the Chat GPT produced answers that fell short of expectations, as it only focused on surface aspects without leading to deeper or critical understanding. This indicates that, although the model was able

to process the instructions, it still struggled to integrate elements of higher-order thinking together in the answers produced (Key 4).

Meanwhile, in key 5, where the questions were related to real cases, the quality produced by both AIs was almost the same. In contrast, the most notable result was seen in the MCQ when using key 6, which is integrating data into Chat GPT. In this case, Chat GPT-4 was able to generate graphs, while Chat GPT-3.5 was unable to present graphs as a representation of data. The graphs in the question allow for the presentation of complex data in a format to facilitate analysis and decision-making, thus supporting students' analytical skills. For key 7, the prompt commands used can generate an outlier answer option. Although the sentences presented are not the same, the purpose of the answer options provided by Chat GPT tends to have a one-way distraction. Overall, each key contributes significantly when applied to Chat GPT in creating HOTS-based questions, especially in the context of entrepreneurship education.

VII. DISCUSSION

In this study, an evaluation was conducted on Chat GPT-4 and Chat GPT-3.5's ability to use 11 strategies to develop HOTS-based assessment questions in the discipline of entrepreneurship we evaluated the ability of Chat GPT-4 and Chat GPT-3.5 using 11 strategies to develop HOTS-based assessment questions in the discipline of entrepreneurship. A series of MCQs were involved on entrepreneurship consisting of entrepreneurship with a total of 807 questions covering the themes of business introduction, marketing strategy, product development, promotion strategy and business management to evaluate comprehensively, broadly and accurately. Both groups of observers assessed each MCQ generated by Chat GPT-4 and Chat GPT-3.5 to obtain diverse views regarding the level of quality and appropriateness of the MCQs to the learning objectives. In particular, Chat GPT-4 demonstrated superior performance in creating HOTS-based MCQs compared to MCQs generated by Chat GPT-3.5 across all indicators, which have been validated by entrepreneurship experts.

MCQs generated by Chat GPT-4 tend to have correct answer choices, but still vary, so that they can function as distractions or traps, especially in HOTS-based MCQ types. In fact, Chat GPT-4 is able to provide correct answers with a high degree of accuracy [11]. However, the main problem found is the presence of distractor answer choices that are obviously wrong or illogical, thus standing out and indirectly revealing the correct answer in the main question. Future research needs to focus on engineering techniques to overcome this problem. Most importantly, 75.9% of all MCQs created by Chat GPT-4 met the set categories.

The superior performance of Chat GPT-4 is in line with expectations and consistent with previous research comparing these two models [43]. This highlights the potential of Chat GPT-4 in understanding and creating human-like conversations to provide answers with rich data, analysis and speed of response. Based on our findings, Chat GPT-4 can generate more relevant questions and answers through an enhanced framework with higher scoring seen in criterion 2. Through this approach, it enables Chat GPT-4 to generate questions and detect more complex and relevant answer options, thus enhancing higher-order cognitive abilities in the right context. In addition, the inclusion of business literature enriched Chat GPT-4's knowledge base.

Furthermore, these findings reveal significant potential when prompt variations are consistently applied to the Chat GPT-4 model, allowing MCQs to encourage critical thinking skills instead of relying solely on memorization. Users are faced with the challenge of analyzing, evaluating, and linking various concepts in answering questions. In addition, the role of Chat GPT-4 in explaining the reasoning process and assessing answers further strengthens its contribution in training deep thinking [44]. The power of Chat GPT-4 in generating questions depends largely on the precision and clarity of the prompt provided. Another study showed similarly, that clear instructions are a key factor in the development of LLM [45]. Therefore, the prompt commands designed for the refinement of LLM instructions become a reinforcement or enhancement of the quality of feedback provided by humans. This design allows LLM to adapt to complicated and uncommon entrepreneurship field HOTS questions. Through specialized training or guidance in applying these 11 keys to Chat GPT, teachers can effectively develop HOTS-based assessment questions.

During our analysis, Chat GPT-4 has potential as a tool in designing questions related to case scenarios and case resolution, where the implied questions can support higher order thinking skills (HOTS). While the final evaluation still requires an observer's judgment, Chat GPT-4's assistance can help teachers in developing more

complex questions. The ability of Chat GPT-4 to generate more precise answers to case-based questions makes it different from the Chat version of GPT-3.5 [46]. Question with case scenarios that have high evaluation scores on learning outcome criteria can be seen in Figure 4.

Interestingly, the entrepreneurship experts consistently gave high ratings to both Chat GPT-4 and Chat GPT-3.5 responses on one criterion. This difference was most noticeable in the assessment of the prompt criteria that addressed views on the strategies used. It is explained about how the prompt can syntactically generate HOTS-based questions, the prompt instructions can logically design questions in accordance with the learning objectives and the clarity of the questions due to the prompt commands given. This proves that the way Chat GPT-4 works is greatly influenced by the prompt instructions received.

The success of Chat GPT in encouraging higher-order thinking has been proven in medical research [47] [48]. Chat GPT's ability to answer higher-order reasoning questions was assessed. In a linear algebra study, Chat GPT was able to create logically sound questions despite requiring evaluation from the instructor [49]. This still shows that Chat GPT can be an effective tool in designing relevant questions. On the other hand, Chat GPT-4 has successfully developed Bloom's Taxonomy questions automatically ranging from lower-order thinking to higher-order thinking specific to Python programming [50]. The study showed that the questions generated by Chat GPT-4 correspond to the elements of Bloom's Taxonomy.

Overall, although Chat GPT-4 showed improvement in responding to complex prompts to generate HOTS-based questions through the use of existing strategies, the technology can still produce bias. Therefore, monitoring the quality of the questions generated is necessary. Without a systematic and in-depth evaluation system, the questions may not meet the desired HOTS standards or may not accurately assess students' competencies. Further research and expert supervision are needed to ensure the applicability of this strategy in education, especially in entrepreneurship subjects at the higher education level.

VIII. IMPLICATIONS

The results of this study have important implications for designing HOTS-based questions for different subject areas and student levels. The suggested strategies provide a framework for designing HOTS-based questions so as to enhance teachers' analytical, critical, creative and digital skills in line with the rapid development of artificial intelligence. The findings show that the strategies integrated in Chat GPT-4 can generate HOTS-based questions compared to Chat GPT-3.5. This is due to the superior ability of Chat GPT-4 to generate more complex and accurate texts. This study can also help teachers in providing question banks flexibly by utilizing Chat GPT, so that they can more easily access and customize questions according to learning needs at various levels of education. In addition, the use of Chat GPT allows the creation of more diverse HOTS-based questions that meet the desired learning objectives.

Despite the great potential of Chat GPT in education, its implementation faces a number of challenges, especially in institutions with limited resources. Infrastructure cost constraints, limited technical skills among educators need to be considered. In addition, issues relating to ethics and data privacy are also concerns that need to be addressed before Chat GPT can be fully integrated into the educational environment. Therefore, it is important to ensure these aspects so that Chat GPT can be used in learning effectively and safely.

Chat GPT is not the only artificial intelligence tool available to generate conversational responses to user questions; there are also artificial intelligence tools such as Bing from Microsoft and Bard from Google. It is hoped that these findings can be utilized to develop the use of Chat GPT in a broader environment, in order to assist the learning process, material development, as well as more effective assessment in various fields of education. Future research could focus on comparing various artificial intelligence platforms in education to explore their advantages and limitations in creating HOTS-based questions.

IX. CONCLUSION

Creating HOTS-based multiple-choice questions has its own challenges, especially to ensure that each question not only tests factual knowledge, but is also able to improve the ability to think critically, analyze, and evaluate concepts in depth in accordance with the principles of Bloom's Taxonomy. To bridge these problems, this study developed 11 strategies applied to Chat GPT to create evaluation questions for entrepreneurship

courses. This method has a detailed discussion in each stage. We believe this approach provides novelty for educators to generate questions in accordance with learning objectives.

Please understand, Chat GPT is a chat box application that requires quick commands to get feedback. Therefore, the importance of clarity of instructions for generating questions on Chat GPT may affect the quality of questions. With this approach, we encourage Chat GPT to be an effective tool to help teachers create entrepreneurial questions that develop higher-order thinking skills in the domain of Bloom's Taxonomy. Furthermore, as a cutting-edge technology, Chat GPT is also considered to have the ability to provide assessment tools comparable to human-generated [51] by adopting tricks and techniques of instruction [52]. Instead of taking a long time, this approach can allow teachers to develop a diverse HOTS-based question bank in a short time using Chat GPT.

With the outstanding capabilities of Chat GPT-4 and the design model, it offers teachers skill enhancement for HOTS-based assessment question generation. Teachers can also focus more on the teaching and evaluation aspects, while the question generation process can be automated and tailored to students' cognitive needs. However, this convenience also poses challenges related to the validity and bias of the questions generated by Chat GPT, especially in the field of entrepreneurship. This issue raises the need for further research to investigate the widespread use of data in the utilization of Chat GPT to develop HOTS-based assessment tools.

Despite these challenges, Chat GPT-4 was able to create questions efficiently, generate varied and complex assessment questions according to the desired aspects and innovate the question structure by verifying the quality, accuracy and relevance of the question scenarios with subject matter experts. Through the developed approach, we hope that educational institutions can consider adopting AI such as Chat GPT to design assessment questions while keeping in mind the rules and ethics of use. In addition, this strategy is also recommended for teachers to use practically in the classroom by combining the use of Chat GPT with traditional assessment methods to improve the quality of questions that will provide actionable insights.

In the future, we encourage further research and collaboration among educators, AI developers and policy makers to improve AI tools in education. Furthermore, there is a need to identify the long-term impact of AI-generated questions on student learning outcomes.

Funding Statement

Funding is supported by Puslapdik Kemendikbudristek, LPDP and BPI for Indonesian education scholarship recipients

Authors Contribution

Almaida Vebibina is the first author who contributed to the conception, review, and writing of the initial draft. Rusijono and Andi Mariono contributed to conception, Ahsan Muafa, Dydik Kurniawan, Ragil Sugeng Dewantoro, Dian Noer Asyari, Sussi Widiastuti, and Jarot Suseno contributed to data collection and processing.

Conflict Of Interest

There is no conflict of interest in writing this article. The authors are responsible for the content of this article.

Acknowledgment

The authors are grateful to Puslapdik Kemendikbudristek, LPDP and BPI for their support and funding. The authors would also like to thank the entrepreneurship experts at Universitas Negeri Surabaya.

REFERENCE

1. Zhou, Y., Gan, L., Chen, J., Wijaya, T. T., & Li, Y. (2023). Development and validation of a higher-order thinking skills assessment scale for pre-service teachers. *Thinking Skills and Creativity*, 48(March), 101272.
2. Maxnun, L., Kristiani, K., & Sulistyningrum, C. D. (2024). Development of HOTS-based cognitive assessment instruments: ADDIE model. *Journal of Education and Learning*, 18(2), 489–498.

3. Häkkinen, P., Järvelä, S., Mäkitalo-Siegl, K., Ahonen, A., Näykki, P., & Valtonen, T. (2017). Preparing teacher-students for twenty-first-century learning practices (PREP 21): A framework for enhancing collaborative problem-solving and strategic learning. *Teaching and Teacher Education: Theory and Practice*, 23(1), 25–41.
4. Tyas, M. A., Nurkamto, J., Marmanto, S., & Laksani, H. (2019). Developing higher-order thinking skills (HOTS)-based questions: Indonesian EFL teachers' challenges. *Journal of Advanced Research in Dynamical and Control Systems*, 2(1), 52–63.
5. Kim How, R. P. T., Zulnaidi, H., & Rahim, S. S. A. (2022). HOTS in quadratic equations: Teaching style preferences and challenges faced by Malaysian teachers. *European Journal of Science and Mathematics Education*, 10(1), 15–33.
6. Ilgun Dibek, M., Sahin Kursad, M., & Erdogan, T. (2024). Influence of artificial intelligence tools on higher-order thinking skills: A meta-analysis. *Interactive Learning Environments*, 1–23.
7. Yalcinalp, S., Türkoğlu, H., Koc, S. E., & Halil, E. (2024). Utilization of artificial intelligence in education: A perspective on learning strategies. *IntechOpen*, 13. Retrieved from <http://dx.doi.org/10.1039/C7RA00172J>.
8. Choi, J. H., Hickman, K. E., Monahan, A. B., & Schwarcz, D. (2022). ChatGPT goes to law school. *Journal of Legal Education*, 71(3), 387–400.
9. Terwiesch, C. (2023). Would Chat GPT get a Wharton MBA? A prediction based on its performance in the operations management course. *Mack Institute for Innovation Management*, 1–26. Retrieved from <https://mackinstitute.wharton.upenn.edu/wp-content/uploads/2023/01/Christian-Terwiesch-Chat-GTP-1.24.pdf>.
10. Eulerich, M., Sanatizadeh, A., Vakilzadeh, H., & Wood, D. A. (2023). Can artificial intelligence pass accounting certification exams? ChatGPT: CPA, CMA, CIA, and EA? *SSRN Electronic Journal*, 1–26.
11. Abeysekera, I. (2024). ChatGPT and academia on accounting assessments. *Journal of Open Innovation: Technology, Market, and Complexity*, 10(1), 100213.
12. Vecchiarini, M., & Somià, T. (2023). Redefining entrepreneurship education in the age of artificial intelligence: An explorative analysis. *International Journal of Management Education*, 21(3), 1–11.
13. Short, C. E., & Short, J. C. (2023). The artificially intelligent entrepreneur: ChatGPT, prompt engineering, and entrepreneurial rhetoric creation. *Journal of Business Venturing Insights*, 19(February), e00388.
14. Bui, H. N., & Duong, C. D. (2024). ChatGPT adoption in entrepreneurship and digital entrepreneurial intention: A moderated mediation model of technostress and digital entrepreneurial self-efficacy. *Entrepreneurship Quarterly*, 19(2).
15. Gupta, P., Mahajan, R., Badhera, U., & Kushwaha, P. S. (2024). Integrating generative AI in management education: A mixed-methods study using social construction of technology theory. *International Journal of Management Education*, 22(3), 101017.
16. Nah, F. F.-H., Zheng, R., Cai, J., Siau, K., & Chen, L. (2023). Generative AI and ChatGPT: Applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research*, 25(2), 277–304.
17. Liu, Y., et al. (2023). Summary of ChatGPT-related research and perspective towards the future of large language models. *Meta-Radiology*, 1(2), 100017.
18. Javaid, M., Haleem, A., Singh, R. P., Khan, S., & Khan, I. H. (2023). Unlocking the opportunities through ChatGPT tool towards ameliorating the education system. *BenchCouncil Transactions on Benchmarks, Standards and Evaluation*, 3(2), 100115.
19. Parker, L., Carter, C., Karakas, A., Loper, A. J., & Sokkar, A. (2024). Graduate instructors navigating the AI frontier: The role of ChatGPT in higher education. *Computers and Education: Open*, 6(July 2023), 100166.
20. George-Reyes, C. E., Vilhunen, E., Avello-Martínez, R., & López-Caudana, E. (2024). Developing scientific entrepreneurship and complex thinking skills: Creating narrative scripts using ChatGPT. *Frontiers in Education*, 9.
21. Chen, L., Ifenthaler, D., & Yau, J. Y.-K. (2021). *Online and blended entrepreneurship education: A systematic review of applied educational technologies* (Vol. 4, No. 2). Springer Singapore.
22. Luan, L., Lin, X., & Li, W. (2023). Exploring the cognitive dynamics of artificial intelligence in the post-COVID-19 and learning 3.0 era: A case study of ChatGPT. *arXiv Preprint arXiv:2302.04818*, 2023, p. 9501.
23. Dwivedi, Y. K., et al. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges, and implications of generative conversational AI for research, practice, and policy. *International Journal of Information Management*, 2023, 71(March).
24. Mavlutova, I., Lesinskas, K., Liogys, M., & Hermanis, J. (2020). Innovative teaching techniques for entrepreneurship education in the era of digitalisation. *WSEAS Transactions on Environment and Development*, 2020, 16, 725–733.
25. Jatmika, R. T. D., Ratnasari, V., & Nadlifatin, R. (2024). Empowering micro-entrepreneurs through artificial intelligence: A conceptual framework for AI-based marketing. *Procedia Computer Science*, 2024, 234, 1087–1094.
26. Chalmers, D., MacKenzie, N. G., & Carter, S. (2021). Artificial intelligence and entrepreneurship: Implications for venture creation in the fourth industrial revolution. *Entrepreneurship Theory and Practice*, 2021, 45(5), 1028–1053.
27. Upadhyay, N., et al. (2022). Influence of digital entrepreneurship and entrepreneurial orientation on the intention of family businesses to adopt artificial intelligence: Examining mediating role of business innovativeness. 2022, 1–65.
28. Xiao, P., Chen, Y., & Bao, W. (2023). Waiting, banning, and embracing: An empirical analysis of adapting policies for generative AI in higher education. *arXiv Preprint arXiv:2305.18617*, 2023, 1–33.
29. Tayan, O., Hassan, A., Khankan, K., & Askool, S. (2024). Considerations for adapting higher education technology courses for AI large language models: A critical review of the impact of ChatGPT. *Machine Learning with Applications*, 2024, 15(May), 100513.

30. Kwangmuang, P., Jarutkamolpong, S., Sangboonraung, W., & Daungtod, S. (2021). The development of learning innovation to enhance higher-order thinking skills for students in Thailand junior high schools. *Heliyon*, 2021, 7(6), e07309.
31. Lewis, A., & Smith, D. (1993). Defining higher-order thinking. *Theory into Practice*, 1993, 32(3), 131–137.
32. Kaminski, J. C., & Hopp, C. (2020). Predicting outcomes in crowdfunding campaigns with textual, visual, and linguistic signals. *Small Business Economics*, 2020, 55(3), 627–649.
33. Prüfer, J., & Prüfer, P. (2020). Data science for entrepreneurship research: Studying demand dynamics for entrepreneurial skills in the Netherlands. *Small Business Economics*, 2020, 55(3), 651–672.
34. Moore, S., Nguyen, H. A., Chen, T., & Stamper, J. (2023). Assessing the quality of multiple-choice questions using GPT-4 and rule-based methods. *Lecture Notes in Computer Science* (including subseries *Lecture Notes in Artificial Intelligence* and *Lecture Notes in Bioinformatics*), 2023, 14200 LNCS, 229–245.
35. Doughty, J., et al. (2024). A comparative study of AI-generated (GPT-4) and human-crafted MCQs in programming education. *Proceedings of the 26th Australasian Computing Education Conference*, 2024, 114–123.
36. Schaper, N. (2024). Using ChatGPT to create constructively aligned assessment tasks and criteria in the context of higher education teaching.
37. Hwang, K., Challagundla, S., Alomair, M. M., Chen, L. K., & Choa, F.-S. (2023). Towards AI-assisted multiple choice question generation and quality evaluation at scale: Aligning with Bloom's taxonomy. 2023. Available at: https://gaied.org/neurips2023/files/17/17_paper.pdf
38. Mellor, R. (2008). *Entrepreneurship for everyone: A student textbook*. SAGE Publications.
39. Anderson, L. W. (2005). Objectives, evaluation, and the improvement of education. *Studies in Educational Evaluation*, 2005, 31(2), 102–113.
40. Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory and Practice*, 2002, 41(4), 212–218.
41. Hopson, M. H., Simms, R. L., & Knezek, G. A. (2001). Using a technology-enriched environment to improve higher-order thinking skills. *Journal of Research on Technology in Education*, 2001, 34(2), 109–119.
42. Zohar, A., & Dori, Y. J. (2003). Higher order thinking skills and low-achieving students: Are they mutually exclusive? *Journal of the Learning Sciences*, 2003, 12(2), 145–181.
43. Eulerich, M., Sanatizadeh, A., Vakilzadeh, H., & Wood, D. A. (2024). Is it all hype? ChatGPT's performance and disruptive potential in the accounting and auditing industries. Springer US, 2024.
44. Ashrafimoghari, V., & Gürkan, N. (2024). Evaluating large language models on the GMAT: Implications for the future of business education. *SSRN Electronic Journal*, 2024, 1–30.
45. Ouyang, L., et al. (2022). Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 2022, 35(NeurIPS).
46. Bendeck, Y., & Voelker, T. A. (2024). ChatGPT's performance on case discussion questions. 2024 (April).
47. Ghosh, A., & Bir, A. (2023). Evaluating ChatGPT's ability to solve higher-order questions on the competency-based medical education curriculum in medical biochemistry. *Cureus*, 2023, 15(4).
48. Sinha, R. K., Roy, A. D., Kumar, N., & Mondal, H. (2023). Applicability of ChatGPT in assisting to solve higher-order problems in pathology. *Cureus*, 2023, 15(2).
49. Lu, K. (2023). Can ChatGPT help college instructors generate high-quality quiz questions? *Human-Computer Interaction and Emerging Technologies (IHET-AI 2023): Artificial Intelligence for Future Applications*, 2023, 70(Ihiet), 311–318.
50. Doughty, J., et al. (2024). A comparative study of AI-generated (GPT-4) and human-crafted MCQs in programming education. 2024. <https://doi.org/10.1145/3636243.3636256>
51. Bhandari, S., Liu, Y., Kwak, Y., & Pardos, Z. A. (2024). Evaluating the psychometric properties of ChatGPT-generated questions. *Computers and Education: Artificial Intelligence*, 2024, 7(August), 100284.
52. Indran, I. R., Paramanathan, P., Gupta, N., & Mustafa, N. (2023). Twelve tips to leverage AI for efficient and effective medical question generation: A guide for educators using ChatGPT. *Medical Teacher*, 2023, 0(0), 1–6.

Appendix 1. Assessment Rubric

Scale section:

A strongly agree: almost all topics agree (SA)

B more agree: many of the topics agree (AM)

C moderately agree: half of the topics agree (MA)

D basically disagree: many topics disagree (BN)

E strongly disagree: almost all topics disagree (SN)

No	Description	Options				
1	Clarity	SA	AM	MA	BN	SN

	Is the question clearly presented and focused so that it can be understood and answered without looking at the options?	A	B	C	D	E
	Are the options longer and contain more detailed information to provide clues to the answer?					
	Are the options longer and contain more detailed information to provide clues to the answer?	A	B	C	D	E
	Does the question allow for multiple correct answer options and require choosing from a range of available combinations?	A	B	C	D	E
	Is there a correct answer among the options provided?	A	B	C	D	E
	Does the MCQ provide enough information to generate answers?	A	B	C	D	E
	Are the answer options different from each other so that the choices provided are unique?	A	B	C	D	E
2	Relevant	SA	AM	MA	BN	SN
	Are the questions and explanations technically correct and error-free??	A	B	C	D	E
	Do the questions and options cover relevant topics and subtopics?	A	B	C	D	E
	Does the question raise any assumptions of controversy?	A	B	C	D	E
	Does the variety of question topics meet the expected criteria or level?	A	B	C	D	E
3	Grammar	SA	AM	MA	BN	SN
	Have the questions and options been framed in clear or unambiguous language?	A	B	C	D	E
	Do the questions and options have consistent grammar?	A	B	C	D	E
	Are the questions and options parallel in style and form?	A	B	C	D	E
4	Prompts	SA	AM	MA	BN	SN
	Is the prompt syntactically correct?	A	B	C	D	E
	Is the prompt logical?	A	B	C	D	E
	Does the prompt support the creation of ambiguity questions?	A	B	C	D	E
5	Goal attainment	SA	AM	MA	BN	SN
	Do the questions describe the type of knowledge required to answer them and the higher order cognitive requirements?	A	B	C	D	E
	Are MCQ answers effective in helping to achieve learning objectives?	A	B	C	D	E
	Do MCQs tend to be negative or fail to measure learning outcomes?	A	B	C	D	E
	Are the questions appropriate for HOTS aspects?	A	B	C	D	E
	Are the questions of good quality?	A	B	C	D	E
	Is the question structure or explanation in line with the learning outcomes?	A	B	C	D	E
	Is the question model and options appropriate for the educational level of the learners?	A	B	C	D	E