Comparing the effectiveness of multiple-answer and single-answer multiple-choice questions in assessing student learning

Yusuf Oc[*], Associate Professor of Marketing, Bayes Business School, City, University of London, UK, yusuf.oc@city.ac.uk

Hela Hassen, Assistant Professor of Marketing, KEDGE Business School, France, hela.hassen@kedgebs.com

*Corresponding author. Address: Bayes Business School (formerly Cass), 106 Bunhill Row,

London, EC1Y 8TZ, United Kingdom

Comparing the effectiveness of multiple-answer and single-answer multiple-choice questions in assessing student learning

Driven by technological innovations, continuous digital expansion has transformed fundamentally the landscape of modern higher education, leading to discussions about evaluation techniques. The emergence of generative artificial intelligence raises questions about reliability and academic honesty regarding multiple-choice assessments in online education. In this context, this study investigates multiple-answer questions (MAQs) versus traditional single-answer questions (SAQs) in online higher-education assessments. A mixed-methods study involving quantitative field experiments and qualitative interviews was conducted with students enrolled in an online Marketing M.Sc. program. The students were divided randomly and assessed using either SAQs or MAQs, and the impacts on test performance of variables such as grade averages, study times, perceived workload, and difficulty were evaluated using independent-sample t-tests and ordinary least-squares regression analysis. The results show that although grades were lower and MAQs were perceived as being more difficult, study times and perceived workload did not differ significantly between the two formats. These findings suggest that despite their challenge, MAQs can promote deeper understanding and greater learning retention. Furthermore, even with their higher perceived difficulty and impact on performance, MAQs hold potential for dealing with academic-integrity concerns related to artificial intelligence.

Keywords: multiple-choice question; online assessment; online quiz, attainment gap, mixed-methods study, field experiment

Introduction

Multiple-choice questions (MCQs) are commonly used for assessments in education, providing a rapid and efficient means of evaluating students' knowledge and understanding. The MCQ format is the dominant one for assessments in online learning (Timmis et al., 2016), and past research has examined the effects of different methods of MCQ development, including sequential, random, chapter contiguous, and reverse test questions (Balch, 1989). When designed properly, MCQs can demand a greater level of analytical thinking, enabling examiners to evaluate students' integration of knowledge, problem-solving skills, and application of knowledge (Rigs et al., 2020, Stevens et al., 2022). Furthermore, electronic assessment using MCQs has several advantages, including objectivity and reliability in assessment, as well as the ability to assess a wide range of skills and provide rapid feedback, which is useful for students (Winstone & Boud, 2022).

However, there are concerns about the limitations of MCQs as assessment tools, especially in online learning, where they can widen attainment and awarding gaps (Cagliesi, 2023; Summers et al., 2023). There are also concerns about the potential for MCQs to rely too heavily on simple, structured problems that assess only factual knowledge and algorithmic solution procedures (Scouller & Prosser, 1994). Such MCQs encourage surface learning and fail to assess higher-order thinking skills such as analysis, synthesis, and evaluation, which are essential for comprehensive understanding and real-world application. Furthermore, the digital era and the proliferation of technological tools have re-engineered contemporary higher education (HE) service offerings, and with the advancement of artificial intelligence (AI), academics are questioning the use of MCQ assessments in their online courses because of the potential for compromised academic integrity (Reddy et al., 2022).

To address the limitations of traditional MCQs, various formats have been used, including true/false questions, single-answer questions (SAQs), multiple-answer questions (MAQs), and scenario-based questions. True/false questions are straightforward but often

oversimplify complex ideas, and SAQs—which require students to select a single correct answer—may not fully capture the depth of a student's understanding (Willson, 1982; Thayn, 2011). By contrast, MAQs—with multiple correct answers—can assess a broader range of knowledge and critical-thinking skills, but they are more challenging for both students and instructors. Meanwhile, scenario-based questions involve applying knowledge to and thinking critically about real-life situations, thereby encouraging deeper learning.

Incorporating MCQs with multiple answers (as used in this study) may encourage students to combine the learning of facts with the understanding of various skills such as analysis, integration, knowledge application, and problem solving (Khoiriyah & Husamah, 2018). A possible side effect of MAQs is that they can increase test-related anxiety more than can other assessments. This is because students must identify several correct answers and avoid incorrect options, which can further heighten anxiety because of the increased complexity and risk of making errors (Jerrim, 2023).

In this study, we focused on two MCQ types: MAQs and SAQs. While both types require students to select a correct answer from a list of options, MAQs allow for the selection of multiple correct answers, whereas SAQs require the selection of a single correct answer. Despite the widespread use of MCQs in educational assessments in marketing education, there is a lack of empirical research comparing the effectiveness of MAQs versus SAQs in evaluating student learning outcomes, particularly at postgraduate level. The aim of this study was to compare the effectiveness of MAQs and SAQs in assessing student learning. By examining the relationship between performance on these two types of MCQs and measures of student achievement, such as exam scores and grades, along with other measures such as study time, perceived workload, perceived difficulty, and course interest, the study sought to determine which MCQ type is more effective in evaluating student understanding and retention of course material.

Previous studies predominantly focused on undergraduate settings, combining MCQs with other assessment types to improve efficacy (Mittal, 2019), addressing issues such as anxiety related to MCQs (Davies, 2015), or comparing their application in different settings, such as computer-based vs. paper-based, or synchronous vs. asynchronous (Downey & Schetzsle, 2012). Hence, there is a need to fully explore the impacts of different MCQ formats on deeper learning and cognitive engagement. The present study addressed this gap with quantitative and qualitative research by investigating not only the performance outcomes but also students' perceptions of difficulty, workload, and overall learning experiences associated with MAQs and SAQs. By doing so, we aim to provide a more nuanced understanding of how different MCQ formats can be leveraged to enhance learning and assessment practices in HE. For marketing educators, we highlight the importance of a combined assessment format involving MAQs and SAQs in HE settings.

Background

Diversifying assessment approaches in higher education: the value of MCQs

Assessments steer the focus of students toward specific subject matter, drive the pace of learning, and direct students' efforts (Scouller & Prosser, 1994). Previous research in different fields has outlined best practices and provided faculty with practical examples for constructing reliable and valid examinations (Gipps, 2005). In HE, various assessment formats are available, such as essay (seen and unseen), *viva voce* (Pearce, 2009), multiple-choice and short-answer questions (Bacon, 2016), and poster session (Stegemman, 2009).

MCQs are used widely for assessments in HE institutions across various countries and disciplines and are particularly prevalent the medical field (DiBattista & Kurzawa, 2011). MCQ examinations that align with educational outcomes can be used to effectively assess various

facets of student performance as well as offering timely feedback and contributing to the process of self-learning (Winston & Boud, 2022). When combined with other methods such as essays or case studies, MCQs can provide a balanced insight into a student's performance, tackling the reliability issues often seen in essays (Bacon, 2003). Online MCQs offer flexibility and choice in assessments, and when integrated with other strategies, they contribute to a thorough assessment approach.

Traditional HE curricula place greater emphasis on knowledge acquisition than on skill development (Janson et al., 2020). This creates an ongoing challenge for educators as they seek to measure students' critical-thinking abilities. Challenges related to online learning are particularly relevant to this study, given that our data comes from online students. Previous research emphasizes issues such as student isolation (Ryan et al., 2001) and the need for active engagement, with multisensory exercises shown to enhance deeper learning (Purinton & Burke, 2020). Hence, educators have long espoused the importance of critical thinking as a means of developing students' higher-order problem-solving abilities (Dahl et al., 2018).

When designed well, MCQs can assess higher-order thinking by targeting analysis, synthesis, and evaluation skills as per Bloom's taxonomy (Bloom, 1956). Such questions prompt students to engage in explaining, analyzing, and evaluating course content, mirroring practical problem-solving situations. However, the effectiveness of MCQs often diminishes when they are constructed poorly; they can focus excessively on knowledge recall and thereby misrepresent students' true understanding (Tarrant et al., 2006).

For MCQs to be effective, careful crafting is essential to encourage inquiry-based learning and to differentiate between levels of student knowledge (Honey & Marshall, 2003). Well-crafted MCQs can assess higher-order cognitive skills beyond mere factual recall, promoting critical thinking, deeper understanding, and learning engagement. For instance, scenario-based questions require students to apply, analyze, and evaluate information, fostering a deeper comprehension of the material (Khoiriyah & Husamah, 2018; Tarrant et al., 2006).

Scenario-based MCQs are a promising avenue, offering real-life problems to which students can apply their knowledge and analytical skills and thus promoting a holistic learning approach (Khoiriyah & Husamah, 2018). The growing emphasis on using MCQs to measure higher-order cognitive skills underscores the importance of crafting questions that deter students from seeking online answers and promote academic rigor (Reddy et al., 2022).

MCQ assessment design: challenges and innovations

An MCQ typically consists of a question (the stem) that identifies the problem at hand, and a set of answer options among which only one is the correct response (the key); within the array of answer options, there are plausible but incorrect responses (the distractors). Constructing MCQ assessments for HE students demands meticulous consideration of various elements, including the question content, format, and structure and the answer options (Scouller & Prosser, 1994; Haladyna & Rodriguez, 2013).

A standard MCQ can be of either an MAQ (students are allowed to select more than one answer) or an SAQ (students are limited to choosing only one response), making them a straightforward and efficient tool for assessing student knowledge (Bacon, 2003). Although SAQs are more common in multiple-choice tests by effectively measuring factual knowledge and recall, they can be inadequate in evaluating complex, higher-order knowledge, skills, and abilities (Thayn, 2011). Researchers have observed that limiting each question to a single correct answer can be restrictive and inadequate for evaluating complex content (Willson, 1982). For example, healthcare professionals often use MAQs to measure intricate, multi-faceted knowledge by requiring examinees to identify multiple components in their responses (van Wijk, 2023). An example from a marketing class might involve analyzing a case study to identify which multiple strategies could improve a company's market position. This approach assesses a student's ability to better reflect the intricacies of real-world problem-solving and the need for comprehensive assessment strategies.

Training in MCQ item writing has expanded, with the aim of bolstering student performance because poor assessment quality can hinder achievement (Tarrant et al., 2006). Poorly crafted items often assess lower cognitive skills (Downing & Yudkowsky, 2005), making MCQ design a labor-intensive task for educators lacking expertise. This challenge potentially leads many educators to source questions from textbook test banks, although such questions are criticized for assessing basic understanding and promoting rote memorizing (Stevens et al., 2022)

AI tools offer significant promise for alleviating the burdens of MCQ design by generating diverse, well-structured questions with plausible distractors tailored to the intended learning outcomes of the course (Rezigalla, 2024). By incorporating generative AI into the MCQ design process, educators can enhance the quality and efficiency of exams, saving considerable time. Nevertheless, human oversight is crucial to ensure the relevance and appropriateness of AI-generated questions because AI may not capture fully the nuances of course content and educational goals.

Navigating test anxiety in MCQ assessments

Test-related anxiety—encompassing test anxiety, computer anxiety, and technostress—is a recognized hurdle in online learning, impacting students during assessments (Davies, 2015; Pekrun, 2019). Marked by fear and uneasiness, test anxiety can hamper exam performance (Reber & Reber, 2001). It often stems from poor preparation, disorganization, or past exam distress, making it difficult for students to grasp and organize course material and hence resulting in negative emotions (Pekrun, 2019).

While some anxiety can sometimes fuel better performance because of perceived workload by increasing study time, participation, and course interest (Naylor, 1997; Paswan & Young, 2002), excessive anxiety levels during tests usually decrease performance (Cassady & Johnson, 2002). However, test performance is not solely about knowing the material; Sweetnam

(2002) noted that a lack of test-taking skills could cause problems for even knowledgeable students. Students might perceive their workload (Estelami, 2015) and the time they spend studying as higher than usual due to test anxiety. Conversely, Igwe and Orluwene (2019) found that employing test-taking strategies could uplift performance and ease exam anxiety.

Formative assessments and thorough guidance before summative assessments have been spotlighted as aids for reducing online test anxiety (Davies, 2015). However, while there are tools in HE that are aimed at enhancing learning experiences (Debuse & Lawley, 2011), challenges remain in managing large classes and ensuring positive experiences for all.

In summary, this comprehensive literature review underscores the crucial role of MCQs in HE assessment, with a specific emphasis on marketing education, and it recognizes the intricate challenges inherent in crafting effective MCQ assessments, particularly for evaluating higher-order cognitive skills. The presence of test anxiety in MCQ assessments is acknowledged as an influential factor shaping student learning and success in HE. Thus, we propose the following hypotheses.

H1: Compared with SAQs, MAQs are associated with (a) lower average grades, (b) increased study time, (c) higher perceived workload, (d) greater perceived difficulty, and (e) less-favorable attitudes toward them.

H2: For both SAQs and MAQs, the final performance in coursework is positively associated with (a) study time, indicating that greater investment in study time enhances student grades, (b) course interest, suggesting that higher engagement with the course material leads to better performance, (c) perceived difficulty, which implies that higher levels of perceived challenge contribute to improved academic outcomes, and (d) perceived workload, indicating that a greater perceived effort is correlated with better coursework performance.

Methodology

Quantitative data

A repeated-measures field experiment was designed to collect data from the same sample of students, who were enrolled in a Marketing Analytics module taught fully online as part of an online M.Sc. degree program in Marketing. Data were collected during July and August 2022: the online module spanned a six-week teaching period followed by two weeks dedicated to assessment and marking, totaling eight weeks. The module comprised primarily asynchronous materials including recorded videos, textual content, readings, and visual examples. Additionally, there were one-hour synchronous online webinars each week where lecturers delved more deeply into concepts, reviewed analyses, and conducted activities such as case-study discussions or online quizzes, and students had the opportunity to ask questions.

At the beginning of the module, the students were divided randomly into two groups based on their weekly activities, and we used these two randomly assigned groups in our experimental design. All the students were given two summative multiple-choice assessments (Test 1 and Test 2), each of which comprised 15 MCQs and contributed 10% of the overall module grade. Two versions of the same test were designed with identical questions but varying answer choices, one version being less challenging and the other more so. We modulated the difficulty level through multiple answer options: a question with multiple correct answers was deemed more challenging than one with a single correct answer. Students received instructions about the tests and question types before they commenced their attempts. Table 1 displays the prompts used prior to the tests and provides a sample question to illustrate the variation in answer choices for the same question.

<< Table 1>>

Both student groups received the same questions; however, in Test 1, the first group encountered classic SAQs whereas the second group was presented with MAQs. Three weeks later, in Test 2, we reversed the question types: the first group tackled MAQs and the second group faced SAQs. After each test, a survey was administered to gather the students' opinions on the effectiveness of the learning process. The students did not know their grades when they were asked to participate in the survey. Additionally, we conducted a focus-group study to gain deeper insights from the students and understand their overall experiences with the two types of multiple-choice assessments.

We expected to see a priming effect (Kahneman, 2011) of our announcement of which group would receive which type of test. The expectation was that upon finding out that they would be taking a more difficult test, those students would spend more time preparing for it.

Data collection

After voluntarily agreeing to participate in the survey and providing their informed consent, the participants responded to a randomized set of questions. These questions addressed (i) the effectiveness of the test on their learning (α=0.863) (1 = reduced my learning; 5 = had a large positive effect on my learning) (Bacon, 2011), (ii) perceived difficulty (α=0.908) (1 = extremely easy; 7 = extremely difficult) (Marks, 2000), (iii) perceived workload (α=0.754) (1 = less than; 2 = equal to; 3 = greater than) (Estelami, 2015), (iv) course interest (α=0.729) (1 = strongly agree; 7 = strongly disagree) (Paswan & Young, 2002). The survey also included other single-scale questions about their perceptions and attitudes, such as "How much time do you think you spent preparing for the quiz?" (1 = less than 3 hours; 5 = more than 12 hours) and their recommendations for the module (1 = strongly recommend; 7 = strongly not recommend) (Peltier, 2003)

The students who took the tests numbered 39 for group 1 and 37 for group 2, totaling 76. We gathered 110 responses from two questionnaires, one administered after each test. After

removing incomplete submissions and responses with missing values, 90 completed surveys (59% response rate) remained: 49 for Test 1 (64% response rate) and 41 for Test 2 (54% response rate). We also recorded the grades that students received in each test and other assessments, in addition to tracking how long they spent on an online learning platform where they could access module materials and view recordings of synchronous sessions. Table 2 summarizes the sample statistics for each test.

<<Table 2>>

Qualitative data

Focus groups are highly recommended for pedagogical research. As suggested by Breen (2006, p. 164), "researchers undertaking pedagogic research might decide to use focus groups when they need to know about student experiences of a particular teaching and/or assessment method". Combined with surveys, focus groups enable researchers to probe participants for more in-depth meanings and produce a more comprehensive description (Wilson, 1997).

We conducted a phenomenological study on postgraduate students enrolled in a marketing program at a business school in Europe. We used online focus-group discussions (FGDs) to investigate their experiences and perceptions of MCQ assessments; to access pertinent information, online students were asked in focus groups to describe their experiences of completing both Test 1 and Test 2. Data were collected through two FGDs, each consisting of five students from the online Marketing M.Sc. program, amounting to a total of 10 students. The discussion guide used in the FGDs is provided in the Appendix. The study was approved by the Research Ethics Committee of the college where the data is collected, and all students were asked to sign a consent form upon their agreement to participate in this study. The researchers emphasized confidentiality, anonymity, and the right to withdraw at the start of each session. To maintain the participants' confidentiality, we gave each of them a pseudonym.

Focus groups are an effective way of gaining qualitative data from discussions when the degree of familiarity with the topic is uniform and when the power relations among discussants are weak and they feel comfortable sharing experiences and exchanging ideas (Creswell, 2012). We assumed that because all the focus-group participants were from the same Marketing Analytics module, they would be more likely to participate actively in the group discussion and disclose their perceptions and experiences of completing the SAQs and MAQs, compared with unacquainted participants from different universities.

Participants, sampling, and recruitment

In this study, we assumed that focus groups comprising 10 international students enrolled in the same module would facilitate the sharing of ideas and concerns with minimal hesitation (Breen, 2006). To capture a variety of perceptions and experiences, we selected five students from Group 1 enrolled in the Marketing Analytics module and five students from Group 2 enrolled in the same module. We expected students from each group to discuss their experiences of MAQs and SAQs. We applied purposive sampling and sent invitation emails to all the students who previously completed the research survey (Creswell, 2012). After receiving a student's agreement via email, researchers made the first contact with the participant and obtained their written consent for research and publication (we disseminated information about the study and its consent forms).

Data collection

In late August 2022, two days after the completion of Test 2, the participants attended the focus-group session. Discussions lasted for 60–90 minutes and were facilitated by a moderator and two academic researchers experienced in conducting focus groups. The participants were divided into two groups (Groups 1 and 2, as previously mentioned). The moderator was informed of his role in accordance with the Practical Guide to Focus-Group Research (Breen, 2006). The moderator made notes to reflect on non-verbal cues.

We conducted FGDs with open-ended and stimulating discussions with the two groups on Microsoft Teams. We started the session by welcoming students in the discussion room and briefing them on the objectives of the study. The participants were asked to give their consent orally before the start of the discussion. Data collection continued until theoretical saturation was reached, which occurred when no new information emerged. This indicates that the data collected were sufficient to capture the full range of the students' experiences and perceptions regarding MAQs and SAQs. All conversations during the session were recorded and transcribed by researchers and research assistants. To guarantee anonymity, participants were not identified. Each participant was given a £20 Amazon voucher for their time.

Data analysis

Transcripts from the FGDs were read and analyzed thoroughly using an inductive approach of thematic analysis (Braun & Clark, 2006) to obtain themes and subthemes. We relied on the three-step formula of Bazeley (2009)—i.e., describe, compare, relate—to work through the results of our analysis. Words and phrases extracted from the transcripts were first individually coded, and agreement among researchers on coding was achieved through Microsoft Teams meetings. We ensured inter-coder reliability by having multiple researchers independently code the data and then discuss discrepancies until consensus was reached. This rigorous process helped to minimize bias and ensured a valid interpretation of the qualitative data.

Our analysis allowed for the extraction of themes and subthemes through the iterative process of describing, comparing, and contrasting the codes within and between the two focus groups. Initially, we identified 15 codes, which were later synthesized into three major themes: 1) Challenging Assessments for Better Learning; 2) Managing Test-Related Anxiety and Time; 3) Positive Impact of MAQs on Learning and Self-esteem. The transition from initial codes to overarching themes was guided by patterns and relationships observed in the data, ensuring that each theme was robust and representative of the participants' experiences. Table 5 reviews the themes and subthemes identified in the FGDs; it illustrates the frequency of each theme's

occurrence in the discussions and provides representative quotes to highlight the students' perspectives.

We used our notes and transcriptions to gain familiarity with the data and master the knowledge of the content in each FGD. We also organized debriefing sessions to avoid missing information that constituted the preliminary analysis. The transcriptions were coded line-by-line using the software package NVivo 11 (QSR International). The coding was conducted by experienced researchers with PhDs and substantial expertise in qualitative research and thematic analysis. As we progressed in the coding process, we identified key themes, including challenges related to MAQs, anxiety and time management, and the promotion of improved learning practices. Our coding process evolved through iterative data analysis, following an inductive, data-driven approach. Codes were grouped into subthemes and overarching themes, reflecting the students' assessment experiences and their learning outcomes.

Results

Comparative analysis of SAQs and MAQs: performance, perceptions, and attitudinal outcomes

We conducted independent-sample t-tests across groups 1 and 2 in both quizzes for variables obtained through the survey and observed variables. The aim was to identify any statistically significant differences between SAQs and MAQs. Table 3 presents the results of the independent-sample t-test analysis.

Consistent with our predictions, tests featuring multiple-answer questions (MAQs) resulted in lower average scores compared to single-answer questions (SAQs), confirming hypothesis H1a. The independent sample t-test showed that students scored significantly higher on tests with SAQs ($M_{test1} = 76.5$, $M_{test2} = 68.4$) than those of tests that contained MAQs ($M_{test1} = 63.7$, $M_{test2} = 59.0$). The difference was statistically significant ($t_{test1} = 3.26$, $p_{test1} < 0.001$; $t_{test2} = 2.04$, $p_{test2} < 0.05$). This result suggests that MAQs pose a greater challenge for students, leading to lower performance outcomes compared to SAQs.

However, hypothesis H1b, which predicted that MAQs would lead to increased study time, was not supported by the data. There was no statistically significant difference in study times between MAQs (M_{test1} = 1112.7, M_{test2} = 1506.2) and SAQs (M_{test1} = 1129.8, M_{test2} = 1587.2), indicating that students spent similar amounts of time preparing for both test formats.

In terms of perceived workload, no significant difference was found between MAQs (M = 3.25) and SAQs (M = 3.19), providing no support for hypothesis H1c. This suggests that students did not perceive MAQs to require more effort or time than SAQs. As for the perceived difficulty, SAQs had a lower mean value (M = 4.42) compared with MAQs (M = 5.74), and the difference was significant ($t = -5.4$, $p < 0.001$); confirming H1d. This finding reinforces the notion that MAQs are viewed as more cognitively demanding.

Looking at the attitudinal differences, it was evident that students held less-favorable opinions of MAQs (M = 4.6) compared with SAQs (M = 5.6), and the difference was statistically significant ($t = 3.2$, $p = 0.002$), confirming hypothesis H1e. Furthermore, students indicated that they would be less likely to recommend MAQs to other students (M_{MAQ} = 5.7, M_{SAQ} = 2.6; $t = -6.9$, $p < 0.001$). We also evaluated the effect of MAQs on students' learning but did not identify any statistically significant influence ($t = 1.6$, $p < 0.061$).

<<Table 3>>

Regression model: effects of independent variables on coursework grades

The second part of our study aimed to examine the impact of several variables on coursework grades, including study time, participation performance, course interest, perceived difficulty of tests, and perceived time spent on coursework. We hypothesized that enhancing the cognitive challenge of tests by including more-complex questions might boost students' engagement with

course material, leading to increased study time and participation performance. Additionally, we predicted that students' interest in the course would be a predictor of their performance in the coursework.

We conducted a regression analysis with coursework grade as the dependent variable, and study time, participation performance, course interest, perceived difficulty of tests, and perceived time spent on coursework as independent variables. The equation for the regression analysis was as follows:

Coursework Grade = β_0 + β_1 Study Time + β_2 Course Interest + β_3 Perceived Difficulty + β_4 Perceived Workload + β_4 Participation Performance + ε.

Where, coursework grade represents the dependent variable, β_0 is the intercept, β1- 4 are the regression coefficients for the respective independent variables, and ε is the error term.

Table 4 gives the results of the regression analysis. The overall model was found to be statistically significant ($F_{(5, 88)}$ = 43.39, $p < 0.001$) and had an R^2 value of 0.72. The analysis revealed that study time on the learning platform (β = 0.254, $p < 0.001$), course interest (β = 1.528, p = 0.006), participation performance (β = 0.178, $p < 0.001$) and perceived difficulty (β = 1.202, p = 0.003) all had significant positive effects on coursework grades, confirming H2a, H2b, and H2c. We did not find any significant influence of perceived workload of the tests on final coursework performance (H2d).

<<Table 4>>

The qualitative data collected from the focus groups were analyzed to identify common themes and subthemes related to students' experiences with MAQs and SAQs. Table 5 below illustrates the frequency of each theme's occurrence in the FGDs and provides representative quotes to highlight the students' perspectives.

<<Table 5>>

17

Qualitative findings derived from our FGDs with students provide valuable insights into the students' perceptions of MAQs and their impact on learning experiences. Students expressed challenges when dealing with MAQs, primarily due to their unfamiliarity with this form of assessment. This aligns with the theme Challenging Assessments for Better Learning, where students perceived MAQs as more difficult and demanding compared to traditional SAQs. They noted that MAQs required them to engage more deeply with the material, promoting their critical-thinking and analytical skills, rather than just encouraging the recall of knowledge. One student remarked: "MAQs were definitely harder than SAQs; I had to think more, for longer and be more focused as well." Another student highlighted how MAQs pushed them to apply concepts more critically: "As I knew that I was going to have MAQs in my test, it required me to really understand and apply the concepts, not just memorise them."

Time management emerged as a significant concern for students when responding to MAQs. The fear of running out of time coupled with the inclusion of scenario-based questions—such as sharing a cluster analysis result table and asking students questions based on that—intensified their anxiety. The increased cognitive load, which stemmed from the need to consider multiple correct answers, further heightened students' anxiety. This is illustrated by the following quote:

It's a new experience for me, good in part and bad in another. Good to be more focused and study harder, but bad in how I lost marks even though I knew almost 90% of the right answers because I was too nervous about ALL the possible correct answers to consider and process and couldn't keep my eyes off the clock. (James)

Despite the inclusion of formative tests with MAQs as part of their module learning, students still reported difficulties in selecting multiple answer options. Our initial survey results indicated that students did not recommend the use of MAQs. However, during the FGDs, a shift in perspective emerged. The focus group discussions provided an interactive environment where

participants could share their experiences, engage in dialogue, and reflect on each other's perspectives. This collective exchange allowed students to express more positive feelings toward MAQs, acknowledging both the challenges they posed and their value in testing deeper learning skills. Through group discussions, students began to perceive MAQs as a fair assessment method that accurately reflected their knowledge. This is illustrated by the following quote:

> This test format compels us to delve deeper and avoid superficial understanding. I firmly believe it provides a fair assessment of my knowledge because it requires more than just remembering. (Karima)

Although there were initial challenges, MAQs were found to benefit students' learning experiences and self-esteem. Students reported that MAQs encouraged them to invest more time in preparation, challenge themselves and deepen their understanding of the course material. This positive impact suggests that MAQs can foster engagement, motivate students to study more thoroughly, and enhance their grasp of course content. Another student claimed:

> I studied more for this test than I have in other modules because I assumed it would be more difficult. I went through each week of content and reviewed it thoroughly. That being said, I found the test reasonably fair and not overly challenging. Although I always aim to excel, I am content with my score, given the circumstances, and proud of myself. (Manon)

Manon's experience underscores how the perceived difficulty of MAQs motivated her to engage more deeply with the course material. The anticipation of a more challenging assessment prompted her to adopt a more rigorous study routine, which involved a comprehensive review of each week's content. This highlights the potential of MAQs to encourage students to invest more time and effort in their studies, thereby enhancing their overall learning experience. The methodical approach taken by Manon, where she reviewed all

course content thoroughly, suggests that MAQs can drive students to develop better study habits. This aligns with our broader theme of Challenging Assessments for Better Learning, indicating that MAQs not only test students' knowledge but also promote a deeper understanding of the material through diligent preparation. Manon's testimony highlights the multifaceted benefits of MAQs in the learning process.

By fostering rigorous study habits, promoting a thorough understanding of the material, ensuring perceived fairness, and enhancing students' self-esteem and motivation, MAQs can serve as effective tools for advancing student learning and academic performance. This analysis demonstrates that while MAQs pose initial challenges, they ultimately contribute to a more engaging and enriching educational experience for students.

Discussion

The need for alternative assessments to classic MCQ tests has been reiterated in previous studies, especially those conducted in the medical field (Sam et al., 2016; van Wijk et al., 2023;) Our study found that MAQs can have numerous benefits for student learning by challenging students but allowing them to succeed, as highlighted by Butler (2018). While our data indicated that students perceived MAQs as more difficult and viewed them less favorably compared with SAQs, we observed no significant differences in the amount of study time or perceived workloads between the two test formats. This suggests that although the preparation time and perceived effort are comparable for the two types of tests, MAQs are still perceived by students as being more challenging and less preferable. Students can develop better study strategies based on their perceptions of the MCQ assessment requiring higher cognitive functioning (Scouller & Prosser, 1994). Previous studies have shown that students' assessment preferences and performance are not correlated (Watering et al., 2008). Congruent with our findings, students exert maximum effort in assessments regardless of their preferences. However, perceived difficulty provides a sharper focus and a sense of accomplishment, particularly when a good score is achieved. We also found that study time on the provided learning platform had a

significant positive impact on coursework grades. In addition, our data showed that course interest and perceived difficulty were positively associated with coursework grades.

To relate our qualitative findings to our survey results, we noted that the theme Challenging Assessments for Better Learning aligned with survey findings indicating that MAQs—despite being perceived as more difficult—were linked to deeper learning and retention. These findings suggest that although MAQs may be more challenging for students, they can lead to higher-order thinking and greater learning outcomes, hence narrowing the awarding gap (Cagliesi et al., 2023). This aligns with previous research on the benefits of challenging and engaging course materials, which have been shown to promote deeper learning and higher levels of student achievement (Stevens et al., 2022). One explanation for these findings is that MAQs may require students to think more critically and engage more deeply with the course material. By requiring students to consider multiple answers and evaluate the relative merits of each option, MAQs may encourage students to go beyond surface-level understanding of the material and develop a more thorough and nuanced understanding. Our regression findings highlight that challenging students in MCQ tests also has an impact on end-of-module assessments, as indicated by Riggs et al. (2020).

While effective for assessing basic knowledge, SAQs may not fully capture the depth of student understanding. Conversely, MAQs, despite their potential to assess higher-order thinking skills, can increase test-related anxiety because of their complexity. Unlike previous studies that suggested the use of SAQs (Simbak et al., 2014), the present study suggests that a combination of SAQs and MAQs is perceived as a fair and challenging evaluation of students' knowledge and skills. As highlighted by Naylor (1997), anxiety motivates students to work harder and thus perform better in exams. Students who took the formative tests were less anxious, had improved attitudes toward tests compared to those who didn't, and acknowledged the benefits of challenging questions, including scenario-based questions that assess higher cognitive skills and encourage deeper learning. Davies (2015) also noted that anxiety associated

with online MCQ tests can be mitigated through the implementation of formative assessments such as practice MCQ tests. To enhance the learning experience associated with MAQs, it is recommended that clear instructions, appropriate time allocation, and support for managing test anxiety should be considered. These adjustments can contribute to a more positive and productive experience with MAQs, thereby benefiting students' learning.

Implications and limitations

Both quantitatively and qualitatively, our study has several limitations that should be considered when interpreting the results. The sample size was small for the experiment, and the study was conducted at only one institution, which may limit the generalizability of our findings. Future research could replicate our study at other institutions or include larger sample sizes to further explore the effects of MAQs on student performance. In addition, when the groups were informed about the different assessments that they would take, both groups were notified simultaneously. This might have influenced both groups to study more, even though they were not all scheduled to take the more difficult test.

We conducted a qualitative study with a diverse group of international distance-learning students. However, we could not eliminate potential language barriers that could affect discussions and interpretation. Focus-group limitations include the emergence of socially acceptable norms and dominant participants influencing the research process.

Nevertheless, despite its limitations, our study illuminates the significant advantages that more-difficult MCQs, such as MAQs, may offer in the realm of student education, similar to previous studies conducted in the medical field (Cerutti, 2019). The insights gained suggest that MAQs could serve as effective pedagogical tools, potentially enhancing students' depth of understanding and elevating their academic success. Therefore, educators should contemplate integrating MAQs into their curriculum to challenge students and stimulate more-comprehensive learning outcomes.

However, it is crucial for educators to recognize that the implementation of MAQs might demand a greater investment of time and effort from students compared with traditional SAQs. Consequently, educators should aim to offer the necessary support and resources, tailoring their teaching strategies to facilitate student success in these more-demanding assessments.

One innovative approach to student assessments—addressing the benefits and potential concerns of the different question types—could be the development of a hybrid MCQ test that combines MAQs and SAQs. This blended format could potentially capitalize on the strengths of both assessment types, using SAQs to gauge baseline comprehension and MAQs to challenge students to apply their knowledge more broadly and critically. Such a balanced approach could mitigate the increased cognitive load associated with MAQs and foster a more nuanced and robust understanding of the course material. It can also help educators address grade inflation issues that occur with SAQ-only quizzes.

To progress with such strategy, it would be beneficial for educators to provide students with formative assessments that mirror this hybrid format, allowing students to adapt to the varied question types by getting immediate feedback for their attempts. Additionally, by offering feedback and guidance on how to navigate MAQs effectively, educators can help students to develop the skills and confidence needed to succeed in these assessments. For example, offering tips on common strategies for answering MAQs or discussing specific pitfalls to avoid can enhance students' ability to navigate these more demanding questions. The goal would be to encourage a more engaged and reflective learning process that prepares students for the complexities of real-world problem-solving.

References

Bacon, D. R. (2003). Assessing learning outcomes: A comparison of multiple-choice and short-answer questions in a marketing context. *Journal of Marketing Education*, *25*(1), 31–36. https://doi.org/10.1177/0273475302250570

Bacon, D. R. (2011). Comparing direct versus indirect measures of the pedagogical effectiveness of team testing. *Journal of Marketing Education*, *33*(3), 348–358. https://doi.org/10.1177/0273475311420243

Bazeley, P. (2009). Analysing qualitative data: More than "identifying themes". *Malaysian Journal of Qualitative Research*, *2*(2), 6–22.

Bloom, B. S. (1956). Taxonomy of educational objectives [Handbook]. *The cognitive domain.* David McKay.

Breen, R. L. (2006). A practical guide to focus-group research. *Journal of Geography in Higher Education*, *30*(3), 463–475. https://doi.org/10.1080/03098260600927575

Butler, A. C. (2018). Multiple-choice testing in education: Are the best practices for assessment also good for learning? *Journal of Applied Research in Memory and Cognition*, *7*(3), 323–331. https://doi.org/10.1016/j.jarmac.2018.07.002

Cagliesi, M. G., Hawkes, D., & Smith, S. (2023). Narrowing awarding gaps: The contributory role of policy and assessment type. *Studies in Higher Education*, *48*(11), 1665–1677. https://doi.org/10.1080/03075079.2023.2209597

Cassady, J. C., & Johnson, R. E. (2002). Cognitive test anxiety and academic performance. *Contemporary Educational Psychology*, *27*(2), 270–295. https://doi.org/10.1006/ceps.2001.1094

Cerutti, B., Stollar, F., Escher, M., Blondon, K., Aujesky, S., Nendaz, M., & Galetto-Lacour, A. (2019). Comparison of long-menu and single-best-answer multiple choice questions in computer-based summative assessments: A randomised controlled trial. *BMC Medical Education*, *19*(1), 219. https://doi.org/10.1186/s12909-019-1651-6

Creswell, J. W. (2012). *Educational research* (4th ed.). Pearson.

Crittenden, W. F., Biel, I. K., & Lovely, W. A. III. (2019). Embracing digitalization: Student learning and new technologies. *Journal of Marketing Education*, *41*(1), 5–14. https://doi.org/10.1177/0273475318820895

Dahl, A. J., Peltier, J. W., & Schibrowsky, J. A. (2018). Critical thinking and reflective learning in the marketing education literature: A historical perspective and future research needs. *Journal of Marketing Education*, *40*(2), 101–116. https://doi.org/10.1177/0273475317752452

Davies, G. (2015). Online MCQ assessment anxiety amongst 1st year psychology students. *Journal of Perspectives in Applied Academic Practice*, *3*(1). https://doi.org/10.14297/jpaap.v3i1.143

Debuse, J. C. W., & Lawley, M. (2011). Using innovative technology to develop sustainable assessment practices in marketing education. *Journal of Marketing Education*, *33*(2), 160–170. https://doi.org/10.1177/0273475311410848

DiBattista, D., & Kurzawa, L. (2011). Examination of the quality of multiple-choice items on classroom tests. *The Canadian Journal for the Scholarship of Teaching and Learning*, *2*(2). https://doi.org/10.5206/cjsotl-rcacea.2011.2.4

Downey, W. S., & Schetzsle, S. (2012). Asynchronous assessment in a large lecture marketing course. *Marketing Education Review*, *22*(2), 97-108.

Downing, S. M. (2005). The effects of violating standard item writing principles on tests and students: The consequences of using flawed test items on achievement examinations in medical education. *Advances in Health Sciences Education: Theory and Practice*, *10*(2), 133–143. https://doi.org/10.1007/s10459-004-4019-5

Gipps, C. V. (2005). What is the role for ICT-based assessment in universities? *Studies in Higher Education*, *30*(2), 171–180. https://doi.org/10.1080/03075070500043176

Haladyna, T. M., & Rodriguez, M. C. (2013). *Developing and validating test items*. Routledge. https://doi.org/10.4324/9780203850381

Igwe, B. M., & Orluwene, G. W. (2019). Test-taking strategies as predictors of students' mathematics achievement in Rivers State. *International Journal of Innovative Psychology & Social Development*, *7*(2), 1–7.

Janson, A., Söllner, M., & Leimeister, J. M. (2020). Ladders for learning: Is scaffolding the key to teaching problem-solving in technology-mediated learning contexts? *Academy of Management Learning and Education*, *19*(4), 439–468. https://doi.org/10.5465/amle.2018.0078

Jerrim, J. (2023). Test anxiety: Is it associated with performance in high-stakes examinations? *Oxford Review of Education*, *49*(3), 321–341. https://doi.org/10.1080/03054985.2022.2079616

Kahneman, D. (2011). Thinking, fast and slow. *Macmillan*.

Mittal, S. (2019). The tyranny of tests: Doing away with tests without selling your soul to grading hell. *Marketing Education Review*, *29*(2), 136-141.

Naylor, F. D. (1997). Test-taking anxiety and expectancy of performance. In J. P. Keeves (Ed.), *Educational research, methodology, and measurement: An international handbook* (2nd ed.) (pp. 971–976). Cambridge University Press.

Paswan, A. K., & Young, J. A. (2002). Student evaluation of instructor: A nomological investigation using structural equation modeling. *Journal of Marketing Education*, *24*(3), 193–202. https://doi.org/10.1177/0273475302238042

Pearce, G., & Lee, G. (2009). Viva voce (oral examination) as an assessment method. *Journal of Marketing Education*, *31*(2), 120–130. https://doi.org/10.1177/0273475309334050

Purinton, E. F., & Burke, M. M. (2020). Engaging online students: Using a multisensory exercise for deeper, active learning. *Marketing Education Review*, *30*(1), 29-42.

Reddy, L., Letswalo, M. L., Sefage, A. P., Kheswa, B. V., Balakrishna, A., Changundega, J. M., Mvelase, M. J., Kheswa, K. A., Majola, S. N. T., Mathe, T., Seakamela, T., & Nemakhavhani, T. E. (2022). Integrity vs. quality of Assessments: Are They Compromised on the Online Platform? *Pedagogical Research*, *7*(2), em0121. https://doi.org/10.29333/pr/11840

Rezigalla, A. A. (2024). AI in medical education: Uses of AI in construction type A MCQs. *BMC Medical Education*, *24*(1), 247. https://doi.org/10.1186/s12909-024-05250-3

Riggs, C. D., Kang, S., & Rennie, O. (2020). Positive impact of multiple-choice question authoring and regular quiz participation on student learning. *CBE—Life sciences Education 19(2)*, Ar.16. https://doi.org/10.1187/cbe.19-09-0189

Ryan, G., Valverde, M., & Rodríguez-Ardura, I. (2001). Marketing education, distance learning and hypermedia: Teaching "current issues in marketing" in a virtual campus. *Marketing Education Review*, *11*(3), 41-53.

Sam, A. H., Hameed, S., Harris, J., & Meeran, K. (2016). Validity of very short answer versus single best answer questions for undergraduate assessment. *BMC Medical Education*, *16*(1), 266. https://doi.org/10.1186/s12909-016-0793-z

Scouller, K. M., & Prosser, M. (1994). Students' experiences in studying for multiple choice

    question examinations. *Studies in Higher Education*, *19*(3), 267–279.

    https://doi.org/10.1080/03075079412331381870

Simbak, N. B., Aung, M. M. T., Ismail, S. B., Jusoh, N. B. M., Ali, T. I., Yassin, W. A. K., . . .

    & Rebuan, H. M. A. (2014). Comparative study of different formats of MCQs: Multiple

    true-false and single best answer test formats, in a new medical school of Malaysia.

    International Medical Journal, *21*(6), 562–566.

Stevens, S. P., Palocsay, S. W., & Novoa, L. J. (2023). Practical guidance for writing multiple-

    choice test questions in introductory analytics courses. *INFORMS Transactions on

    Education*, *24*(1), 51–69. https://doi.org/10.1287/ited.2022.0274

Summers, R. J., Burgess, A. P., Higson, H. E., & Moores, E. (2023). How you teach and who

    you teach both matter: Lessons from learning analytics data. *Studies in Higher

    Education*, 1–16.

Sweetnam, K. R. (2002). *Test taking strategies and student achievement running head*.

Tarrant, M., Knierim, A., Hayes, S. K., & Ware, J. (2006). The frequency of item writing flaws

    in multiple-choice questions used in high stakes nursing assessments. *Nurse Education

    in Practice*, *6*(6), 354–363. https://doi.org/10.1016/j.nepr.2006.07.002

Thayn, S. (2011). *An evaluation of multiple choice test questions deliberately designed to

    include multiple correct answers*. Brigham Young University. Retrieved from

    https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=3449&context=etd

Timmis, S., Broadfoot, P., Sutherland, R., & Oldfield, A. (2016). Rethinking assessment in a

    digital age: Opportunities, challenges and risks. *British Educational Research Journal*,

    42(3), 454–476. https://doi.org/10.1002/berj.3215

Van de Watering, G., Gijbels, D., Dochy, F., & Van der Rijt, J. (2008). *Students' assessment preferences, perceptions of assessment and their relationships*.

van Wijk, E. V., Janse, R. J., Ruijter, B. N., Rohling, J. H. T., van der Kraan, J., Crobach, S., Jonge, M., Beaufort, A. J., Dekker, F. W., & Langers, A. M. J. (2023). Use of very short answer questions compared to multiple choice questions in undergraduate medical students: An external validation study. *PLOS ONE*, *18*(7), e0288558. https://doi.org/10.1371/journal.pone.0288558

Willson, V. L. (1982). Maximizing reliability in multiple-choice questions. *Educational and Psychological Measurement*, *42*(1), 69–72. https://doi.org/10.1177/0013164482421008

Wilson, V. (1997). Focus groups: A useful qualitative method for educational research? *British Educational Research Journal*, *23*(2), 209–224. https://doi.org/10.1080/0141192970230207

Winstone, N. E., & Boud, D. (2022). The need to disentangle assessment and feedback in higher education. *Studies in Higher Education*, *47*(3), 656–667. https://doi.org/10.1080/03075079.2020.1779687

Appendix

Focus Group Discussion Guide

Introductory questions:

What do you think of the programme?

Tell me about the module Marketing Analytics, how was it compared to the other modules?

Transitioning questions:

Tell me about your assessments for the module Marketing Analytics and more specifically regarding the 2 tests that you recently completed.

- If you think back about **TEST 1**, how did you feel about it? because you knew you were going to have certain type of test- did you prepare more/ or less than usual (compared to your previous modules)?

General Perceptions and Experiences

Initial Impressions: What were your initial thoughts and feelings when you first encountered MAQs in your assessments?

I went through the survey results and noticed that for Test 1 some of the comments suggested that the quiz was fair, and similar to other MCQ you had in the past.

Is that a general feeling – do you all think that?

Why did you respond so? Could you please elaborate on this?

Now let's focus on **TEST 2**, that you completed this week, how did test 2 go?

Was it more challenging or less challenging?

Would you say you have **prepared equally** for test 1 and test 2? If not, why? (Did you spend more time studying for Test 2? Test 1?)

Comparative Views: How do you compare MAQs to other forms of assessment, such as Single Answer Questions (SAQs) or essay-type questions? In what ways do you find MAQs more or less challenging than other assessment types?

Which tests reflects best your knowledge and why?

Effectiveness and fairness: Do you believe MAQs are a fair way to assess your knowledge and understanding of the course material? Why or why not?

How effective do you think MAQs are in testing your comprehensive understanding of the subject?

Do you believe SAQs are a fair way to assess your knowledge and understanding of the course material? Why or why not?

How effective do you think SAQs are in testing your comprehensive understanding of the subject?

Study Habits: Can you describe any changes in the way you approached your study material as a result of MAQs?

Did knowing that MAQs would be part of your assessment influence your study habits and preparation methods?

Specific Challenges and Benefits

Time Management: How did you manage your time during Test 1 and Test 2?

Were there any specific strategies you used to ensure you completed the tests within the allotted time?

Cognitive Load and Anxiety: What specific challenges did you face when answering SAQs and MAQs, and how did you address them? Did you experience any anxiety or stress related to SAQs and MAQs? If so, can you elaborate on these experiences?

Critical Thinking and Understanding: How did SAQs and MAQs affect your engagement with the course material and your critical thinking skills? Can you provide examples of how a question encouraged you to think more deeply or critically about the subject matter?

Personal reflections and suggestions:

Overall, does the way assessment works here help you to learn well, or does it interfere or cause you difficulties?

What improvements, if any, would you suggest for the implementation of SAQs and MAQs in future assessments?

Are there any specific aspects of SAQs and MAQs that you found particularly beneficial or detrimental to your learning experience?

How do you feel SAQs and MAQs have influenced your overall learning experience and academic performance?

Long-term Impact: Do you think the skills and knowledge you developed while preparing for SAQs and MAQs will benefit you in future courses or professional scenarios? If so, how?