Ardura, D., Zamora, A., Calderón-Lima, R. & Galán, A. (2026). Question Type Matters: Capturing Students' Learning Approaches in Educational Assessment. *Revista de Investigación Educativa*, 44. https://doi.org/10.6018/rie.642511

Question Type Matters: Capturing Students' Learning Approaches in Educational Assessment

El tipo de pregunta importa: capturando los enfoques de aprendizaje del alumnado en la evaluación educativa

Diego Ardura[1]*, Ángela Zamora**, Rebeca Calderón-Lima*** y Arturo Galán*

*Departamento de Métodos de Investigación y Diagnóstico en Educación I. Universidad Nacional de Educación a Distancia (España)

** Departamento de Métodos de Investigación y Diagnóstico en Educación II. Universidad Nacional de Educación a Distancia (España)

***Investigadora independiente, Formación del Profesorado y Ciencias del Aprendizaje (USA)

Abstract

Different types of questions are used to assess students' knowledge, and it is widely recognized that students' learning approaches significantly influence their performance on these questions. In this study, we investigated how different types of questions capture students' learning approaches. The sample consisted of 140 secondary school students. Descriptive, correlational and path analyses were implemented. Results showed that the surface motive approach was negatively related to academic performance in open-ended questions, but not in multiple-choice questions. Moreover, analyses revealed a mediating effect of academic self-efficacy between learning approaches and academic performance when measured with problem-solving questions. However, no indirect effects were found when academic achievement was assessed with multiple-choice or short-answer questions. Open-ended questions were able to capture students' learning approaches more effectively than close-ended questions. Given the relationship between question types and students' approaches to learning, pedagogical implications for the design of effective assessments are discussed.

Keywords: Academic achievement, educational testing, approaches to learning, question types

1 **Correspondence:** Diego Ardura, dardura@edu.uned.es, c/ Juan del Rosal, 14. 28040, Madrid.

Resumen

Para evaluar los conocimientos del alumnado se utilizan diferentes tipos de preguntas, y existen evidencias de que los enfoques de aprendizaje influyen significativamente en su rendimiento en estas preguntas. En este estudio, se investiga cómo los diferentes tipos de preguntas captan los enfoques de aprendizaje de las y los alumnos. La muestra estuvo compuesta por 140 alumnos y alumnas de secundaria. Se utilizaron análisis descriptivos, correlacionales y de senderos. Los resultados mostraron que el enfoque superficial se relacionó negativamente con el rendimiento académico en las preguntas abiertas, pero no en las preguntas de opción múltiple. Además, los análisis revelaron un efecto mediador de la autoeficacia académica entre los enfoques de aprendizaje y el rendimiento académico cuando se medía con preguntas de resolución de problemas. Sin embargo, no se encontraron efectos indirectos cuando se evaluaba el rendimiento académico con preguntas de opción múltiple o de respuesta corta. Las preguntas abiertas permitieron captar los enfoques de aprendizaje del alumnado de forma más eficaz que las preguntas cerradas. Dada la relación entre los tipos de preguntas y los enfoques de aprendizaje de las y los alumnos, se discuten las implicaciones pedagógicas para el diseño de evaluaciones eficaces.

Palabras clave: Rendimiento académico; pruebas de evaluación; enfoques del aprendizaje, tipos de preguntas.

Introduction

The notion of academic achievement is a pervasive concept in the field of education (York et al., 2019). As a result, the measurement of academic success and the study of the factors that influence students' performance have been broadly examined in recent decades (see, for example, Alyahyan & Düştegör, 2020). At each stage of education, different methods are used to collect information about students' learning and performance, with written examinations being one of the most commonly used assessment tools. In turn, these exams can be designed using a wide variety of question types. The types of questions range from close-ended to open-ended, depending on whether students are given a predefined set of possible answers. The first group typically includes true/false (TFQ) and multiple-choice (MCQ) questions, while fill-in-the-blank, short-answer (SAQ), essay and problem-solving questions (PSQ) constitute the open-ended or constructed-response alternatives (Simkin & Kuechler, 2005). The choice of the appropriate question type, not only to properly assess students' knowledge but also to stimulate the learning process, has been the subject of significant debate between researchers and practitioners (see for example Baburajan et al., 2020).

The use of close-ended questions has increased significantly over the last few decades, primarily due to practical reasons that have made them one of the most popular question types. For example, they are more efficient than open-ended questions, especially when administering tests to large groups of students (Snyder, 2003). They also allow teachers to cover a wide range of course content (Becker & Johnston, 1999). Moreover, they are perceived as more objective (Becker & Johnston, 1999) and facilitate immediate feedback after examinations (Delgado & Prieto, 2003). From the

student's perspective, close-ended questions do not require a deep understanding of the content, especially when the students perceive that the test only involves memorizing facts (Biggs, 1987). Besides, the likelihood of selecting the correct answer increases as students can often narrow down their options by eliminating the least plausible options from the given set of answers (Bush, 2001). Furthermore, it has been demonstrated that reverse objective testing enhances higher-order reasoning skills more effectively than traditional objective testing (Martínez-Abad et al., 2024).

However, research comparing the use of open-ended and closed-ended questions has yielded inconclusive results. Early studies mainly compared MCQs with different types of SAQs. While some researchers reported that student performance was similar in both formats (Bleske-Rechek et al., 2007; Schladitz et al., 2017), others concluded that MCQs are a better assessment tool than open-ended questions. For example, Bridgeman and Lewis (1994) reported that, although they were both correlated, MCQs were a better predictor of students' GPA than open-ended questions (OEQs). In fact, Bridgeman and Morgan (1996) even suggested that MCQs and OEQs may measure different cognitive abilities.

More recent research has raised concerns about the limitations of MCQs and suggests the use of SAQs, or at least a combination of both, for a more comprehensive assessment of student knowledge (Radad et al., 2023; Stanger-Hall, 2012). For example, a comparison of MCQs and open-ended questions in an astronomy course revealed that MCQs may overestimate students' knowledge, as many students were unable to give correct justifications for their answers (Wooten et al., 2014). The overestimation of students' performance using close-ended questions was also reported in recent studies (Breuer et al., 2023; Couch et al., 2018). In contrast, SAQ seemed to be able to capture students' knowledge more effectively. In the same study, the authors suggested that none of the question types alone is satisfactory for measuring student learning. A combination of question types seems to be necessary for comprehensive student assessment. This aligns with findings from various fields such as mathematics problem solving, where Thacker et al. (2013) found that a large percentage of students who chose the correct answer to a MCQ were unable to provide an adequate justification for their choice. In creativity research, open-ended questions have been associated with more creative individuals (Acar et al., 2023). In medical research, a recent study found that very SAQs (i.e., open-ended) were more reliable and discriminating than multiple-choice questions (van Wijk et al., 2023). In fact, guidelines for designing MCQs suggest that distractors should be based on previously analysed open-ended questions (Schwarz, 2023). In addition, university teachers indicated that MCQs can be used to assess higher-order cognition, but only at the "apply" or "analyse" levels (Liu et al., 2024; Xiromeriti & Newton, 2024). Furthermore, a previous study showed the effectiveness of SAQs in retrieval practice for university students, whereas MCQs do not produce these positive effects (Greving & Richter, 2022).

Extensive research has been reported on the variables associated with academic achievement (Gutiérrez-de-Rozas et al., 2022). Among all the predictors of academic performance, psychological traits such as students' learning approaches (Cairns & Areepattamannil, 2022) and self-efficacy (Schneider & Preckel, 2017) have been

shown to be key in explaining student performance. Depending on their learning goals, students can approach their learning with different approaches (Biggs, 1987). Two approaches to learning (AL) have previously been proposed: deep and surface (Biggs et al., 2001). Each approach is based on the students' motives to learn and the strategies they follow during learning. Students with a deep AL are intrinsically motivated and focused on understanding the meaning of what they are studying, whereas those with a surface approach perceive learning tasks as a requirement to be met and their motivation is more likely to be extrinsic. Previous research has found an association between AL and academic achievement. In a meta-analysis, Schneider and Preckel (2017) reported a significant negative relationship between surface approaches and academic performance. Other studies confirmed that surface learning approaches were significant negative predictors of performance in primary school mathematics (García et al., 2016) and secondary school science (Rosander & Bäckström, 2014). In a previous study, Ardura and Galán (2019) found that self-efficacy was not only a predictor of academic achievement, but also an effective mediator variable between learning approaches and academic performance in secondary school students. More recently, Laitinen et al. (2024) found that students with low self-efficacy tended to adopt superficial learning strategies in the context of computer science education.

As noted above, previous studies on the use of different question types have mainly focused on direct comparisons between them, rather than investigating the effects of relevant psychological constructs such as ASE and AL. Only a few studies have examined the effect of AL on performance using different question types. In educational research, Scouller (1998) showed that university students with a deep AL tended to perform worse on MCQs than those with a surface approach. Conversely, the trend reversed when essays were used as the assessment tool, with deep learners outperforming surface learners. Furthermore, Yonker (2011) claimed that the negative impact of surface AL on MCQ performance in higher education was greater than the positive effect of deep AL on the same question type. Similarly, in the context of a professional bachelor's degree, Baeten et al. (2008) found that surface AL was a negative predictor of academic performance when the assessment was conducted using a portfolio. Taken altogether, these studies seem to show that the choice of specific question types appears to influence students' adoption of particular AL and their performance.

The aim of the study

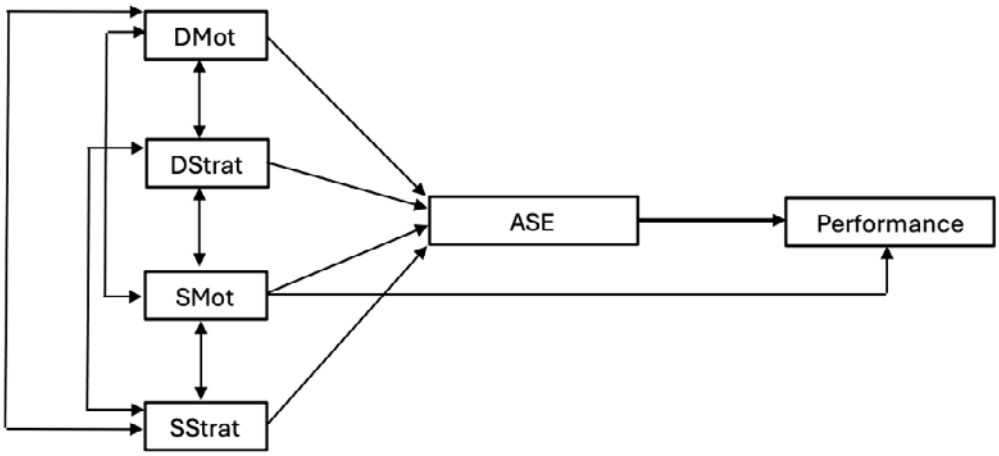
As highlighted earlier, teachers usually employ different types of questions to measure students' academic performance. However, there is a paucity in the research that compares the effects of learning approaches and self-efficacy on the students' performance when assessed through close-ended and open-ended questions. Furthermore, previous studies have focused mainly on university students, limiting our understanding of these effects in other student populations.

Given the relevance of academic self-efficacy and learning approaches in student academic performance, the aim of the present study was to analyse how these factors

influence secondary school students' outcomes across three commonly used question types in written examinations: SAQs, MCQs, and PSQs. Thus, our research was guided by two research questions: RQ1. How does secondary students' performance on different types of questions reflect their learning approaches? and RQ2. Does students' academic self-efficacy mediate the relationship between their learning approaches and their performance on close-ended and open-ended questions? To answer these questions, a path analysis was conducted for each type of question (SAQs, MCQs, and PSQs), building on findings from a previous study that explored the effects of learning approaches and self-efficacy on secondary students' performance in a science subject (Ardura & Galán, 2019) (see Figure 1).

Figure 1

Tentative model for the effect of students' academic self-efficacy and learning approaches on academic achievement



Note. ASE: Academic self-efficacy; DMot: deep approach to learning motive; DStrat: deep approach to learning strategies; SMot: surface approach to learning motive; SStrat: surface approach to learning strategies

Method

Participants

This research was conducted within the context of a third-year chemistry course in Spain's compulsory secondary education system. The sample was selected using convenience sampling based on accessibility. A total of 140 secondary school students (45.0% male and 55.0% female) participated in the study. Average age of participants was 14.4 years.

Instruments

For this investigation, student academic performance across three different types of questions served as the criterion variables, while students' learning approaches and academic self-efficacy were used as the predictor variables.

- *Academic performance across different types of questions*. To measure student academic performance on different question types, a test was designed. The test was developed following the Spanish framework for secondary school learning standards (ALS), which outlines the competencies students are expected to achieve each school year (Ley Orgánica 8/2013, de 9 de diciembre, para la mejora de la calidad educativa). This test was part of the class assessment and therefore had an impact on students' grades. The test contained three different types of questions. The first type of question included one SAQ focused on conceptual topics, such as defining terms or explaining scientific laws. Examples include: 'List the steps of the scientific method' or 'State Boyle's law and provide a graphical interpretation'. The second type of question consisted of a set of three multiple-choice items. These items were designed to assess both conceptual knowledge ('It is NOT a fundamental quantity in the International System of Units') and application questions ('If the real value of a quantity is 4.00 cm and we measure 4.50 cm, the relative error would be'). Finally, the third question included one numerical problem; for example: 'A gas occupies a volume of 300mL at a temperature of 30oC. Assuming the pressure remains constant, what temperature change is required for the gas to occupy a volume of 1L? Performance on each type of question was measured on a scale of 0 to 10 points. In addition, an assessment rubric for the test was designed by one of the authors in collaboration with a science teacher. This instrument included several indicators for assessing the open-ended questions, each aligned with corresponding achievement levels. The students' tests were independently and anonymously marked by one of the authors and a secondary school teacher. Any discrepancies in marking were discussed later until agreement was reached.
- *Academic self-efficacy*. Students' academic self-efficacy (ASE) was measured using the self-efficacy for learning and performance subscale included in the Spanish version of the Motivated Strategies for Learning Questionnaire (MSLQ) (Pintrich et al., 1991). This instrument has previously been successfully adapted and validated in a sample of Spanish students (Albert Pérez, 2017). The subscale consists of 8 items, such as 'I am confident that I can understand the basic concepts taught in this course'. Participants are asked to indicate their level of agreement with each item on a 7-point Likert scale, where 1 indicates they 'strongly disagree' and 7 indicates they 'strongly agree'. The validity of the instrument was established by CFA (c2/df=1.24, GFI=.91, CFI=.86, RMSEA=.041). In the present study, the reliability of this subscale, measured by McDonald's Omega, reached a value of 0.73.
- *Learning approaches*. The Spanish adaptation of the Revised Two-Factor Study Process Questionnaire (R-SPQ-2F) developed by Biggs et al. (2001) was used

to assess students' learning approaches. This instrument was validated for the Spanish population of secondary school students by Blanco et al. (2009) and de la Fuente et al. (2008). Moreover, this instrument has recently been used in research involving secondary school students (Skorbakk & Gamlem, 2025). The original questionnaire consists of 20 items. Ten items make up the deep approach scale, and the other ten items comprise the surface approach scale. The deep approach is characterized by an active search for understanding. Learners who adopt a deep approach are motivated by intrinsic interest and a genuine engagement with their learning process. Conversely, the surface approach is characterized by a more mechanical, compliance-oriented learning style, usually motivated by external factors such as the desire to pass exams or avoid failure. Each scale has two subscales, motives and strategies, and each of them contains 5 items. Motives refer to the underlying reasons why a learner engages in the learning process, whereas strategies refer to the specific techniques or methods a learner uses to study and learn content. For example, the item, 'I find that studying sometimes gives me a sense of deep personal satisfaction,' would measure the deep motive approach to learning; whereas the item, 'I find that I can get through most exams by memorizing key sections rather than trying to understand them,' would measure the surface strategy approach to learning. The CFA showed reasonably good fit indices for a four-factor model: $\chi^2/df=2.20$, GFI=.93, CFI=.89 and RMSEA=.068. A reliability analysis using McDonald's Omega showed values between .67 and .72 for the four subscales.

Procedure

Prior to the study, the research project was presented to the school board. After their approval, students and families were contacted to explain the objectives of the investigation and how the data was going to be gathered and protected. Both the students and their families gave their informed consent, which was obtained through forms gathered prior to data collection. Confidentiality and anonymity policies were enforced. One of the authors designed the chemistry test given to students, carried out the data gathering and was responsible for marking the exam together with another secondary school teacher. Data collection was conducted in the classroom. First, students completed a 10-minute survey that included the instruments to measure self-efficacy and learning approaches. Second, the students were given 40 minutes to complete the chemistry test. To ensure anonymity, each student was assigned a code provided by the school.

Data Analyses

All variables followed a normal distribution with kurtosis and skewness values between -1 and +1. Confirmatory factor analysis (CFA) was used to validate the instruments, and reliability was assessed using the McDonald's Omega statistic. Pearson coefficients were calculated to estimate correlations between variables. To obtain explanatory models of how learning styles and self-efficacy influence student perfor-

mance on each type of question, a separate path analysis model using the maximum likelihood method was developed for each question type. As it is known, the goodness of a model, which is the degree of agreement between the experimental data and the theoretical model, can be measured using several indices proposed in the literature (Blunch, 2013). The most demanding criterion and the one used in this study is the χ^2 test, which tests the null hypothesis of agreement between the model and the data. Goodness of fit (GFI), comparative goodness of fit (CFI) and root mean square error approximation (RMSEA) are also reported: $1 < \chi^2/df < 3$; CFI > .90; GFI > .90; RMSEA < .08 (Schumacker & Lomax, 2016). Finally, 95% bias-corrected bootstrap confidence intervals based on 5000 bootstrap samples was used for the investigation of mediation effects (Hayes, 2018). SPSS (v.29.0) and AMOS (v.19.0) were used to perform the calculations (IBM Corp., 2020).

Results

In this section, we present the results of our path analyses, which were conducted to answer the study's research questions. To strengthen the contextual description, we will first introduce the findings of the descriptive and correlational analyses of the study variables.

Descriptive and correlational analyses

Table 1 shows the means and standard deviations of all the variables included in the study. Performance was higher on multiple-choice questions (MCQ) with a mean of 7.26, while performance on open-ended questions (SAQ) and problems (PSQ) had lower means: 6.60 and 5.75, respectively. In terms of learning approaches, the means for deep learning motives (DMot) and strategies (DStrat) were 3.14 and 2.96 accordingly. Mean scores for surface motives (SMot) and surface strategies (SStrat) were 2.42 and 3.01 respectively.

A positive correlation was found between student performance across the different types of questions. The strongest relationship was found between the two open-ended questions (SAQ) and problem-solving tasks (PSQ), with a correlation coefficient of .431 ($p < .01$). Furthermore, ASE correlated positively with performance in the case of SAQ ($r = .250$, $p < .01$) and PSQ ($r = .324$, $p < .01$), but no significant relationship was found when academic performance was measured using MCQ (see Table 1). The relationships between the different learning approaches and student performance on each type of question were mixed. On the one hand, no significant relationships were found between MCQ performance and any of the four different ALs. On the other hand, in the case of the SAQ and PSQ, negative significant correlations were found between those question types and the surface learning subscales. More specifically, SAQ and PSQ were both negatively correlated to the surface approach to learning motivation ($r = -.354$, $p < .01$ for SAQ and $r = -.306$, $p < .01$ for PSQ) and the surface approach to learning strategies ($r = -.195$, $p < .01$ for SAQ and $r = -.154$, $p < .01$ for PSQ). Conversely, only the SAQ was positively correlated to the deep learning approach motive ($r = .140$, $p < .05$, see Table 1).

Table 1

Descriptive statistics and correlations

	M	SD	01	02	03	04	05	06	07	08
01. SAQ	6.60	3.60	1							
02. MCQ	7.26	2.72	.173*	1						
03. PSQ	5.75	3.33	.431**	.132	1					
04. ASE	5.05	0.81	.250**	.031	.324**	1				
05. DMot	3.14	0.71	.147*	.004	.046	.336**	1			
06. DStrat	2.96	0.66	.098	.032	.023	.331**	.615**	1		
07. SMot	2.42	0.79	-.347**	.008	-.327**	-.468**	-.423**	-.311**	1	
08. SStrat	3.01	0.77	-.354**	-.030	-.306**	-.324**	-.359**	-.221**	.720**	1

Note: SAQ: academic achievement in open-ended questions; MCQ: academic achievement in multiple-choice questions; PSQ: academic achievement in problems; ASE: Academic self-efficacy; DMot: deep approach to learning motive; DStrat: deep approach to learning strategies; SMot: surface approach to learning motive; SStrat: surface approach to learning strategies. * $p < .05$; ** $p < .01$

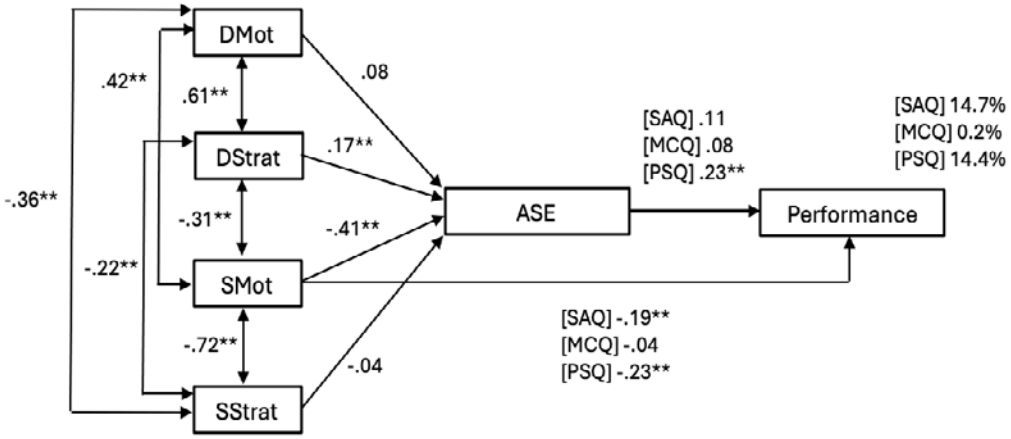
Path analyses

A path analysis for each type of question was developed. To answer the first research question, we examined the direct relationships between learning approaches and academic performance in each of the question types. Only the surface motive approach showed a statistically significant correlation with student performance in the SAQ and PSQ. The proposed model was then tested to explore the second research question which examines the mediating effect of academic self-efficacy between learning approaches and students' performance in each question type (see Figure 2). Thus, the four AL variables were exogenous in the model, whereas ASE and student performance in each question type were endogenous variables. Following this general scheme, a structural equation model was generated for each question type.

The three models presented a good adjustment to data. Model 1 (SAQ): (df=3), p=.413, GFI=.998, CFI=.999, RMSEA=.002; Model 2 (MCQ): (df=3), p=.928, GFI=.999, CFI=.997, RMSEA=.003; Model 3 (PSQ): (df=3), p=.427, GFI=.993, CFI=.998, RMSEA=.01. Besides, the models explained a higher percentage of variance in open-ended questions (SAQ and PSQ), 14.7% and 14.4% respectively, in comparison to the MCQ which showed only 0.2%.

Figure 2

Path analyses for the different types of questions



Note. SAQ: academic achievement in open-ended questions; MCQ: academic achievement in multiple-choice questions; PSQ: academic achievement in problems; ASE: Academic self-efficacy; DMot: deep approach to learning motive; DStrat: deep approach to learning strategies; SMot: surface approach to learning motive; SStrat: surface approach to learning strategies. * $p < .05$; ** $p < .01$

In the models, the surface motive approach was directly and negatively related to academic performance in open-ended questions ($\beta = -.23$, $p<.01$ for PSQ and $\beta = -.19$, $p<.01$ for SAQ). However, in the case of the multiple-choice questions this relationship was non-significant (see Figure 2). Deep strategy and surface motive approaches to learning were positively ($\beta = .17$, $p<.01$) and negatively ($\beta = -.41$, $p<.01$) associated with academic self-efficacy, respectively. In turn, academic self-efficacy was correlated to performance, only when this trait was measured by problem solving.

Table 2 summarizes the mediation analyses of academic self-efficacy between learning approaches on question performance. Statistically significant mediation effects between learning approaches and academic performance through self-efficacy were found only in the case of performance on the problem-solving question where two indirect effects were present. The deep strategy ($\beta = .19$, $p=.034$) and surface motive approaches ($\beta = -.38$, $p=.006$) showed an indirect effect on performance. The remaining performance measures (i.e. SAQs and MCQs) showed no indirect effects of academic self-efficacy and learning approaches.

Table 2

Test for mediation using a bootstrap analysis with a 95% confidence interval.

Relationships	Direct effects	Indirect effects	Confidence intervals		p
			Low	High	
Short answer question					
DStrat→ASE→ SAQ	>.01	.08	-.04	.35	.167
DMot→ASE→ SAQ	>.01	.03	-.04	.24	.283
SStrat→ASE→ SAQ	>.01	.01	-.08	.21	.524
SMot→ASE→ SAQ	-1.32	-.16	-.53	.11	.200
Multiple Choice Question					
DStrat→ASE→MCQ	>.01	.03	-.05	.20	.377
DMot→ASE→ MCQ	>.01	.01	-.03	.16	.373
SStrat→ASE→ MCQ	>.01	>.01	-.04	.14	.471
SMot→ASE→ MCQ	.10	-.06	-.31	.15	.464
Problem Solving Question					
DStrat→ASE→ PSQ	>.01	.19	.01	.51	.034
DMot→ASE→ PSQ	>.01	.08	-.09	.34	.278
SStrat→ASE→ PSQ	>.01	.03	-.24	.29	.696
SMot→ASE→ PSQ	-.95	-.38	-.81	-.11	.006

Note. SAQ: academic achievement in open ended questions; MCQ: academic achievement in multiple choice questions; PSQ: academic achievement in problems; ASE: Academic self-efficacy; DMot: deep approach to learning motive; DStrat: deep approach to learning strategies; SMot: surface approach to learning motive; SStrat: surface approach to learning strategies. Note: Unstandardized coefficients reported. Bootstrap sample: 5000 with replacement.

Discussion and conclusions

Previous studies have highlighted the key role of designing teaching strategies that foster deep learning approaches among students (Chin & Brown, 2000). Assessment plays a critical role in teaching practices, significantly shaping how students approach their studies and, consequently, their learning outcomes (Furnham et al., 2011). In fact, previous research has shown that the type of questions included in tests can strongly influence students' approaches to learning (Furnham et al., 2011). This study aimed to examine how different types of questions, commonly employed across educational settings, reflect and capture students' learning approaches. In addition, a mediator effect of academic self-efficacy between students' learning approaches and performance on different types of questions was explored.

Results showed that structural equation models related performance on PSQs to learning approaches through two different mechanisms. First, there was a direct, negative relationship between the surface motive approach and performance on

PSQs. In addition, two indirect effects were observed via ASE: one linking the deep strategy approach and the other linking the surface motive approach to performance on PSQs. In the case of open-ended questions, the surface motive approach had a direct effect on academic performance, but this effect was specific to this type of assessment. In contrast, when MCQs were used to measure performance, no significant relationships- direct or indirect- were found. These findings align with Scouller's (1998) study, which reported that students tended to use a surface learning approach when they anticipated MCQs as the format of a future test. In other words, these results highlight that neither deep nor surface learning approaches are significantly associated with performance on MCQs. Consequently, a student's choice to rely less on specific learning approaches seems justified, as these strategies may not guarantee a higher score in this type of closed-ended test. Additionally, these findings indicate that open-ended questions may be a more sensitive assessment tool to examine students' knowledge, as their performance on such questions appears to be more closely tied to their learning approaches.

An alternative explanation for the absence of direct or indirect effects of learning approaches on MCQ performance may lie in the strategies students use to answer such questions. For instance, Funk and Dickson (2011) reported a similar effect in psychology students, suggesting that students may use different mechanisms when answering MCQs. For example, they may be able to recognize the correct answer to a question or, without knowing the correct answer, discard incorrect answers to deduce the right choice. These strategies are not available to students when they must independently generate answers to open-ended questions. Similarly, Wooten et al. (2014) found that a high percentage of astronomy undergraduates were unable to justify their correct choice in a MCQ item. Our results suggest that students may not need to use deep learning approaches to successfully answer MCQs, as test-taking strategies alone might be enough to get correct answers. This finding may explain the previously reported prevalence of students' subject-matter misconceptions, which appear to persist when success on MCQs relies more on test-taking strategies than on deep conceptual understanding (Couch et al., 2018). On the other hand, this phenomenon might be related to the kind of construct each question type assesses. For example, performance on open-ended questions tends to correlate more strongly with itself than with close-ended questions, indicating that the two formats may measure different underlying constructs. This observation supports previous suggestions that open-ended and close-ended questions might evaluate distinct aspects of student knowledge and cognitive processes (Bridgeman, 1992; Hubbard et al., 2017; Thacker et al., 2013).

Our path analyses also showed that the predictive power of ASE varies depending on the type of question. In particular, ASE was a better predictor of performance on open-ended questions compared to closed-ended questions. Considering that self-efficacy influences how people allocate effort and time to complete a task (Beck & Schmidt, 2018), the results of this study suggest that more cognitively demanding tasks, such as PSQ, require greater mastery of their self-regulatory skill than tasks like MCQs, where students are not required to generate an answer independently. This finding aligns with previous findings suggesting that students tend to perceive close-ended questions as less challenging than open-ended questions (Biggs, 1987; Bush, 2001). In addition,

it is important to note that the short open-ended questions used in this research were declarative in nature, directly related to the content being assessed. In this context, self-efficacy was also found to be a non-significant predictor of performance. This suggests that the perceived complexity of a question may play a crucial factor in determining the extent to which self-efficacy influences academic performance.

It is worth emphasizing the need to minimize superficial approaches to learning, as their avoidance appears to play a more significant role in academic achievement than adopting a deep learning approach, particularly in explaining performance in chemistry. This observation has also been highlighted in recent research within the context of primary mathematics education (García et al., 2016), and in a recent meta-analysis (Schneider & Preckel, 2017).

Educational implications

The findings of this study highlight several key considerations for educators in designing effective assessment strategies. First, it is important for teachers to be aware that students may develop 'testwiseness' skills, as demonstrated by DeVore et al. (2016) in their investigation of the effect of test-taking strategies in physics assessments. In other words, students may develop skills which help them to improve their performance in tests, regardless of their actual knowledge of the subject. To mitigate this, teachers should design assessments that prioritize the evaluation of students' actual understanding rather than test-taking strategies. Moreover, it is well known that the type of questions included in a test is a key factor in how students approach and direct their learning (Furnham et al., 2011). Undoubtedly, as discussed above, closed-ended questions have many practical advantages, especially in terms of grading efficiency. However, results in this study suggest that closed-ended questions may not effectively differentiate between students who engage in deep versus superficial learning approaches. Open-ended questions, or constructed-response questions, appear to capture the learning approaches of students who try to understand the material deeply. Generating answers to such questions can encourage students to use higher-order thinking skills during their study time and facilitate deep learning. This echoes previous studies that have found that some students' approaches to learning are linked to critical thinking (Laird et al., 2014). In secondary school chemistry, including open-ended questions can help students engage with the material at a deeper level, as such questions may lead students to apply conceptual and procedural knowledge. Nevertheless, some contexts such as massive courses make the use of open-ended questions more difficult. In these cases, and, in view of our results, alternatives to conventional objective tests should be considered if students' deep learning approaches are meant to be reflected in test assessments. A remarkable alternative has been recently proposed by Martínez-Abad et al. (2024). These authors have shown how reverse objective testing promotes high-order thinking skills, which have been shown to be related to deep learning approaches (Zhang & Sternberg, 2000).

Encouraging students to use deep learning approaches has been identified as an effective strategy for improving student performance (Bunce et al., 2017). Our models further suggest an interesting indirect effect of ASE, which mediates the relationship

between deep learning approaches and academic performance, particularly in the context of constructed-response questions. Given the importance of deep learning approaches in meaningful understanding and achievement, our findings suggest that teachers should design lessons not only to encourage the adoption of this type of approach but also to actively foster students' self-efficacy. This means that by building confidence in their abilities, educators can stimulate students' deep engagement with the material, ultimately leading to better learning outcomes.

Limitations and prospective

While this study provides valuable insights, limitations must be acknowledged to contextualize the findings. The first limitation of this study lies in the measurement of approaches to learning and academic self-efficacy. Although the use of self-report measures is common in educational research, relying on students' self-assessment may introduce some bias into the results. Future studies could use additional measures that do not require self-reported data and perform data triangulation to obtain more objective measurements. The second limitation relates to the fact that only a limited number of questions of each type were answered by the students in our research. This may restrict the robustness of the findings, so future research should replicate these results with a larger and more diverse set of questions. Additionally, future replication studies should examine the effect of student age on the observed effects. Given that this study was conducted within the context of a science subject, further studies should explore whether similar patterns hold across different academic domains. Finally, based on our results, future studies should explore any other potential limitations of different types of questions in terms of capturing students' learning approaches.

References

Acar, S., Berthiaume, K., & Johnson, R. (2023). What kind of questions do creative people ask? *Journal of Creativity*, *33*(3), 100062. https://doi.org/10.1016/j.yjoc.2023.100062

Albert Pérez, A. (2017). *Evaluación del aprendizaje autorregulado: validación del Motivated Strategies Learning Questionnaire en educación secundaria* [Tesis doctoral, Universitat de València]. Repositori d'Objectes Digitals per a l'Ensenyament, la Recerca i la Cultura. http://hdl.handle.net/10550/59163

Alyahyan, E., & Düştegör, D. (2020). Predicting academic success in higher education: literature review and best practices. *International Journal of Educational Technology in Higher Education*, *17*(1), 3. https://doi.org/10.1186/s41239-020-0177-7

Ardura, D. & Galán, A. (2019). The interplay of learning approaches and self-efficacy in secondary school students' academic achievement in science. *International Journal of Science Education, 41*(13)1723–1743. https://doi.org/10.1080/09500693.2019.1638981

Baburajan, V., e Silva, J. D. A., & Pereira, F. C. (2020). Open-ended versus closed-ended responses: A comparison study using topic modeling and factor analysis. *IEEE transactions on intelligent transportation systems*, *22*(4), 2123-2132. https://doi.org/10.1109/TITS.2020.3040904

Baeten, M., Dochy, F., & Struyven, K. (2008). Students' approaches to learning and assessment preferences in a portfolio-based learning environment. *Instructional Science*, *36*, 359-374. https://doi.org/10.1007/s11251-008-9060-y

Beck, J. W., & Schmidt, A. M. (2018). Negative Relationships Between Self-Efficacy and Performance Can Be Adaptive: The Mediating Role of Resource Allocation. *Journal of Management*, *44*(2), 555-588. https://doi.org/10.1177/0149206314567778

Becker, W. E., & Johnston, C. (1999). The Relationship between Multiple Choice and Essay Response Questions in Assessing Economics Understanding. *Economic Record*, *75*(4), 348-357. https://doi.org/10.1111/j.1475-4932.1999.tb02571.x

Biggs, J. (1987). *Student Approaches to Learning and Studying*. Australian Council for Educational Research.

Biggs, J., Kember, D., & Leung, D. Y. P. (2001). The revised two-factor Study Process Questionnaire: R-SPQ-2F. *British Journal of Educational Psychology*, *71*(1), 133-149. https://doi.org/10.1348/000709901158433

Blanco, A., Prieto, L., Torre, J.C., & García, M. (2009). Adaptación, validación y evaluación de la invarianza factorial del cuestionario revisado de procesos de estudio (R-SPQ-2F). En A. Boza (coord.), *Actas del IX Congreso Nacional de Modelos de Investigación Educativa sobre 'Educación, investigación y desarrollo social'* (pp. 1535–1543). Universidad de Huelva.

Bleske-Rechek, A., Zeug, N., & Webb, R. M. (2007). Discrepant performance on multiple-choice and short answer assessments and the relation of performance to general scholastic aptitude. *Assessment & Evaluation in Higher Education*, *32*(2), 89-105. https://doi.org/10.1080/02602930600800763

Blunch, N. J. (2013). *Introduction to Structural Equation Modeling Using IBM SPSS Statistics and AMOS* (2nd ed.). Sage.

Breuer, S., Scherndl, T., & Ortner, T. M. (2023). Effects of response format on achievement and aptitude assessment results: multi-level random effects meta-analyses. *Royal Society Open Science*, *10*(5), 220456. https://doi.org/10.1098/rsos.220456

Bridgeman, B. (1992). A Comparison of Quantitative Questions in Open-Ended and Multiple-Choice Formats. *Journal of Educational Measurement*, *29*(3), 253-271. https://doi.org/10.1111/j.1745-3984.1992.tb00377.x

Bridgeman, B., & Lewis, C. (1994). The Relationship of Essay and Multiple-Choice Scores With Grades in College Courses. *Journal of Educational Measurement*, *31*(1), 37-50. https://doi.org/10.1111/j.1745-3984.1994.tb00433.x

Bridgeman, B., & Morgan, R. (1996). Success in College for Students with Discrepancies between Performance on Multiple-Choice and Essay Tests. *Journal of Educational Psychology*, *88*(2), 333-340. https://doi.org/10.1037/0022-0663.88.2.333

Bunce, D. M., Komperda, R., Schroeder, M. J., Dillner, D. K., Lin, S., Teichert, M. A., & Hartman, J. R. (2017). Differential Use of Study Approaches by Students of Dif-

ferent Achievement Levels. *Journal of Chemical Education*, *94*(10), 1415-1424. https://doi.org/10.1021/acs.jchemed.7b00202

Bush, M. (2001). A Multiple Choice Test that Rewards Partial Knowledge. *Journal of Further and Higher Education*, *25*(2), 157-163. https://doi.org/10.1080/03098770120050828

Cairns, D., & Areepattamannil, S. (2022). Teacher-Directed Learning Approaches and Science Achievement: Investigating the Importance of Instructional Explanations in Australian Schools. *Research in Science Education, 52*, 1171–1185 https://doi.org/10.1007/s11165-021-10002-0

Chin, C., & Brown, D. E. (2000). Learning in Science: A Comparison of Deep and Surface Approaches. *Journal of Research in Science Teaching*, *37*(2), 109-138. https://doi.org/10.1002/(SICI)1098-2736(200002)37:2<109::AID-TEA3>3.0.CO;2-7

Couch, B. A., Hubbard, J. K., & Brassil, C. E. (2018). Multiple–true–false questions reveal the limits of the multiple–choice format for detecting students with incomplete understandings. *BioScience*, *68*(6), 455-463. https://doi.org/10.1093/biosci/biy037

de la Fuente, J., Pichardo, M. C., Justicia, F., & Berbén, A. (2008). Enfoques de aprendizaje, autorregulación y rendimiento en tres universidades europeas / Learning approaches, self-regulation and achievement in three European universities. *Psicothema*, *20*(4), 705-711. https://www.psicothema.com/pdf/3544.pdf

Delgado, A. R., & Prieto, G. (2003). The effect of item feedback on multiple-choice test responses. *British Journal of Psychology*, *94*(1), 73-85. https://doi.org/10.1348/000712603762842110

DeVore, S., Stewart, J., & Stewart, G. (2016). Examining the effects of testwiseness in conceptual physics evaluations. *Physical Review Physics Education Research, 12*(2), 020138. https://doi.org/10.1103/PhysRevPhysEducRes.12.020138

Funk, S. C., & Dickson, K. L. (2011). Multiple-Choice and Short-Answer Exam Performance in a College Classroom. *Teaching of Psychology*, *38*(4), 273-277. https://doi.org/10.1177/0098628311421329

Furnham, A., Batey, M., & Martin, N. (2011). How would you like to be evaluated? The correlates of students' preferences for assessment methods. *Personality and Individual Differences*, *50*(2), 259-263. https://doi.org/10.1016/j.paid.2010.09.040

García, T., Rodríguez, C., Betts, L., Areces, D., & González-Castro, P. (2016). How affective-motivational variables and approaches to learning predict mathematics achievement in upper elementary levels. *Learning and Individual Differences*, *49*, 25-31. https://doi.org/10.1016/j.lindif.2016.05.021

Greving, S., & Richter, T. (2022). Practicing retrieval in university teaching: short-answer questions are beneficial, whereas multiple-choice questions are not. *Journal of Cognitive Psychology*, *34*(5), 657-674. https://doi.org/10.1080/20445911.2022.2085281

Gutiérrez-de-Rozas, B., López-Martín, E., & Carpintero Molina, E. (2022). Condicionantes del rendimiento académico: revisión sistemática de 25 años de meta-análisis. *Revista de Educación*, *398*, 39–85. https://doi.org/10.4438/1988-592X-RE-2022-398-552

Hayes, A. F. (2018). *Introduction to Mediation, Moderation, and Conditional Process Analysis a Regression-Based Approach* (2nd ed.). Guilford Press.

Hubbard, J. K., Potts, M. A., & Couch, B. A. (2017). How Question Types Reveal Student Thinking: An Experimental Comparison of Multiple-True-False and Free-Response Formats. *CBE-Life Sciences Education*, *16*(2), ar26. https://doi.org/10.1187/cbe.16-12-0339

IBM Corp. (2020). *IBM SPSS Statistics for Windows* (27.0) [Computer software]. IBM Corp.

Laird, T. F. N., Seifert, T. A., Pascarella, E. T., Mayhew, M. J., & Blaich, C. F. (2014). Deeply Affecting First-Year Students' Thinking: Deep Approaches to Learning and Three Dimensions of Cognitive Development. *The Journal of Higher Education*, *85*(3), 402-432. https://doi.org/10.1080/00221546.2014.11777333

Laitinen S., Christopoulos, A., Laitinen, P., & Nieminen, V. (2024). Relationships between self-efficacy and learning approaches as perceived by computer science students. *Frontiers in Educ*ation, *9*, 1181616. https://doi.org/10.3389/feduc.2024.1181616

Ley Orgánica 8/2013, de 9 de diciembre, para la mejora de la calidad educativa. *Boletín Oficial del Estado*, *295*, de 10 de diciembre de 2013. https://www.boe.es/eli/es/lo/2013/12/09/8/con

Liu, Q., Wald, N., Daskon, C., & Harland, T. (2024). Multiple-choice questions (MCQs) for higher-order cognition: Perspectives of university teachers. *Innovations in Education and Teaching International*, *61*(4), 802-814. https://doi.org/10.1080/14703297.2023.2222715

Martínez-Abad, F., Hernández-Ramos, J. P., Sánchez-Prieto, J. C., Izquerdo-Álvarez, V., del Moral Marcos, M. T., Rivetta, M. S., & Ortíz-López, A. (2024). ¿Innovar en el examen tipo test? La prueba objetiva inversa para mejorar la evaluación sumativa en educación superior. *REDU: Revista de Docencia Universitaria, 22*(2), 233-250. https://doi.org/10.4995/redu.2024.21752

Pintrich, P. R., Smith, D. A. F., García, T., & McKeachie, W. J. (1991). *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)*. National Center for Research to Improve Postsecondary Teaching and Learning.

Radad, K., Taha, M., & Rausch, W. D. (2023). Multiple Choice Questions Versus Very Short Answered Questions in the Evaluation of Students of Veterinary Pathology. *Revista Española de Educación Médica*, *4*(1), 27-35. http://doi.org/10.6018/edumed.548861

Rosander, P., & Bäckström, M. (2014). Personality traits measured at baseline can predict academic performance in upper secondary school three years late. *Scandinavian Journal of Psychology*, *55*(6), 611-618. https://doi.org/10.1111/sjop.12165

Schladitz, S., Ophoff, J., & Wirtz, M. (2017). Effects of different response formats in measuring Educational Research Literacy. *Journal for Educational Research Online*, *9*(2), 137-155. https://doi.org/10.25656/01:14900

Schneider, M., & Preckel, F. (2017). Variables associated with achievement in higher education: A systematic review of meta-analyses. *Psychological Bulletin*, *143*(6), 565-600. https://doi.org/10.1037/bul0000098

Schumacker, R. & Lomax, R. G. (2016). *A beginner's guide to structural equation modelling* (4th ed.). Routledge.

Schwarz, G. (2023). Multiple-Choice Questions for Teaching Quantitative Instrumental Element Analysis: A Follow-Up. *Journal of Chemical Education*, *100*(10), 4099-4105. https://doi.org/10.1021/acs.jchemed.3c00061

Scouller, K. (1998). The Influence of Assessment Method on Students' Learning Approaches: Multiple Choice Question Examination versus Assignment Essay. *Higher Education*, *35*(4), 453-472.

Simkin, M. G., & Kuechler, W. L. (2005). Multiple-Choice Tests and Student Understanding: What Is the Connection? *Decision Sciences Journal of Innovative Education*, *3*(1), 73-98. https://doi.org/10.1111/j.1540-4609.2005.00053.x

Skorbakk, I., & Gamlem, S. M. (2025). Exploring self-assessment practices and learning approaches in science among upper secondary students. *Assessment in Education: Principles, Policy & Practice, 32*(3), 299-319. https://doi.org/10.1111/j.1540-4609.2005.00053.x

Snyder, A. (2003). The New CPA Exam-Meeting Today's Challenges: The Revised Exam Simplifies the Process Not Only for Those Taking the Test but Also for Their Employers. *Journal of Accountancy*, *196*(6), 11.

Stanger-Hall, K. F. (2012). Multiple-choice exams: An obstacle for higher-level thinking in introductory science classes. *CBE—Life Sciences Education, 11*(3), 294–306. https://doi.org/10.1187/cbe.11-11-0100

Thacker, B., Chapagain, G., Pattillo, D., & West, K. (2013). The Effect of Problem Format on Students' Responses. https://doi.org/10.48550/arXiv.1312.6004

van Wijk, E. V., Janse, R. J., Ruijter, B. N., Rohling, J. H., van der Kraan, J., Crobach, S., de Jonge, de Beaufort, A. J, Dekker, F.W., & Langers, A. M. (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

Wooten, M. M., Cool, A. M., Prather, E. E., & Tanner, K. D. (2014). Comparison of performance on multiple-choice questions and open-ended questions in an introductory astronomy laboratory. *Physical Review Special Topics - Physics Education Research*, *10*(2). https://doi.org/10.1103/PhysRevSTPER.10.02010

Xiromeriti, M., & Newton, P. M. (2024). Solving not answering. Validation of guidance for writing higher-order multiple-choice questions in medical science education. *Medical Science Educator, 34*(6), 1469-1477. https://doi.org/10.1007/s40670-024-02140-7

Yonker, J. E. (2011). The relationship of deep and surface study approaches on factual and applied test-bank multiple-choice question performance. *Assessment & Evaluation in Higher Education*, *36*(6), 673-686. https://doi.org/10.1080/02602938.2010.481041

York, T. T., Gibson, C., & Rankin, S. (2019). Defining and measuring academic success. *Practical Assessment, Research, and Evaluation*, *20*(1), 5. https://doi.org/10.7275/hz5x-tx03

Zhang, L. F., & Sternberg, R. J. (2000). Are learning approaches and thinking styles related? A study in two Chinese populations. *The Journal of psychology, 134*(5), 469-489. https://doi.org/10.1080/00223980009598230

Fecha de recepción: 17 diciembre, 2024
Fecha de revisión: 6 enero, 2025
Fecha de aceptación: 9 diciembre, 2025