International Journal of Evaluation and Research in Education (IJERE)
Vol. 15, No. 4, August 2026, pp. 2946~2958
ISSN: 2252-8822, DOI: 10.11591/ijere.v15i4.38745

Motivational and normative drivers of generative AI substitution in academic work: a mixed-methods study from Saudi higher education

Mazin Mansory, Zilal Meccawy
English Language Institute, King Abdulaziz University, Jeddah, Saudi Arabia

Article Info

Article history:

Received Jan 17, 2026
Revised May 21, 2026
Accepted Jun 26, 2026

Keywords:

Academic integrity
Generative AI
Institutional clarity
Moral rationalization
PLS-SEM
Saudi higher education

ABSTRACT

The rapid integration of generative artificial intelligence (GenAI) in higher education has intensified tensions between legitimate learning support and unauthorized task substitution, particularly where institutional guidance remains ambiguous. This mixed-methods study investigates how attitudes toward AI, moral rationalization strategies, and perceived institutional clarity interact to shape AI-based substitution behavior among 249 undergraduates at a Saudi university. Partial least squares structural equation modeling (PLS-SEM) revealed that positive attitudes and rationalization together explained 46% of the variance in substitution behavior, with perceived clarity of institutional guidance significantly moderating the rationalization–substitution link. Complementary interviews with seven students and ten instructors revealed that linguistic burden, peer norms, and fragmented faculty guidance facilitated boundary crossing from scaffolding to shortcutting. By exploring the relationship between moral neutralization and environmental clarity, this research offers a new approach to evaluating the effectiveness of institutional AI guidance beyond common technology acceptance models. The findings can be used to inform the design of multi-tiered, inclusive AI usage policies, assessments that value process over product, and culturally responsive academic integrity education within a multilingual higher education context.

Corresponding Author:

Mazin Mansory
English Language Institute, King Abdulaziz University
P.O. Box 80200, Jeddah 21589, Saudi Arabia
Email: mmmansory@kau.edu.sa

1. INTRODUCTION

The normative framework of academic integrity has long been established in higher education, fostering the development of ethical, self-regulating learners within a community built on ethical and responsible trust [1]. Generative artificial intelligence (GenAI) has revolutionized this paradigm, offering not just access to information but ready-to-use text, fully written, sourced, cited, and reflected on, as though it were an original, fully elaborated scholarly text [2]. While prior digital tools (such as the digital aid) had separate digital support structures corresponding to different stages of the academic process, GenAI consolidates multiple stages (and functions) within the complete academic process (idea generation, brainstorming, drafting, editing, revision, and proofreading) to one location, which creates a blurred line of distinction between scaffolding (a legitimate form of 'pedagogical assistance') and shortcuts (an unethical way of 'substituting' work) [3]. The majority of institutional responses to academic dishonesty with the use of AI tools focus on controlling through punitive policies or technologies. Most institutional responses to cheating using AI tools are through behavioral controls (e.g., banning AI writing, using detection technology,

and punitive policy language) [4]. Unfortunately, most detection systems are not 100% accurate; therefore, they often incorrectly classify non-native English speakers' (NNESs) writing as AI-generated content (rather than as original) [5]. In addition, by taking a prohibitionist approach, institutions fail to help students develop the self-regulatory and moral reasoning skills necessary for successfully navigating technology-enhanced learning environments [6], [7].

These trends most certainly apply to the fast-changing landscape of higher education within the Kingdom of Saudi Arabia, since the launch of Vision 2030 with digital transformation and AI preparedness being of national interest [8], [9]. There has been a major investment by universities in infrastructural improvements relating to integrity, such as: honor codes, the use of plagiarism detection software, and the implementation of proctoring systems (all being geared towards the previous understanding of what constitutes academic misconduct, before GenAI) [10]. It appears as though that the current rulings around how to regulate the use of feedback loops from humans to machines are insufficient for preventing the use of a machine to assist in developing any of the various tasks involved with producing a written work (e.g., drafting, constructing an argument, and creating a synthesis). In cases where NNES are learning in a second language and are more likely to be engaged in high-stakes assessments, the use of GenAI serves a dual purpose for NNES: it offers a legitimate means of providing access to language that has "been polished," and it serves as an easy escape from completing the cognitive processes that are required to complete high-stakes assessments (the purpose of which is to foster the cognitive processes) [11], [12]. While global discourse grows, data from Saudi higher education on how students and educators perceive the contours of appropriate AI use are lacking.

This research attempts to bridge this gap by synthesizing a motivational view (theory of planned behavior or TPB) with a norm-based view (neutralization theory) of how attitude, rationalization, and perceived institutional clarity influence AI-based substitution behavior. The contributions of our research are threefold. Crucially, these contributions extend existing scholarship in three ways. First, the study repositions neutralization theory as a predictive mechanism embedded within a TPB-grounded model, and introduces institutional clarity as an empirically testable moderator of the rationalization-to-behavior pathway; together, these moves offer a moral-cognitive analytical framework for AI-mediated academic boundary crossing that complements rather than replaces existing technology acceptance approaches and, to the best of our knowledge, has not been previously applied to multilingual higher education contexts. Second, by employing neutralization theory not as an independent explanatory model but as a moderator of student attitude and substitutive behavior, the framework foregrounds a moral-cognitive mechanism that is underexplored in technology acceptance theories. Third, by positing institutional clarity as a contextual moderator constraining rationalization, the study contributes initial empirical evidence of 'integrity de-risking through transparency' in a Saudi university setting. Finally, the sequential explanatory mixed-methods design triangulates student motivations with pedagogical challenges, demonstrating integration beyond parallel data collection through joint display meta-inferences [13], [14].

2. LITERATURE REVIEW

2.1. GenAI and the scaffold–shortcut boundary

Academic integrity is a complex phenomenon where individual judgement, social context, assessment, and culture of the institution collide [15], [16]. GenAI tools such as ChatGPT are unique to the degree that they actively reclassify technology from a tool for accessing information to a tool for generating academic outputs, whenever written text, argumentation, paraphrased products, and analytic tasks are needed on demand [2]. Many students utilize AI in certain situations when they are attempting to gain clarity about a concept or to translate. Students will also utilize AI in direct production of content relating to their assignment (for example, by creating text) [17]. AI usage that is legitimate in one course may come across as an offence in another course, particularly in situations where the instructor does not provide concrete or clear examples of what AI practices are considered acceptable and unacceptable [18].

The scaffold/shortcut boundary is defined as the transition point between using AI as a support tool in enhancing a student's cognitive process and replacing that process entirely with AI. In other words, it is the differences between when AI provides the essential elements (key ideas, structure of evidence) for the student to use in developing their own intellectual material versus when the student simply submits the work as their own without any cognitive effort [3]. The distinction is the foundation for this research project, and will be referred to as "the boundary" throughout this paper. While AI-based substitution involves replacing certain steps in the process of creating an assignment with the help of a computer program (or other technological tools), contract cheating requires the student to have someone else do the assignment for them, often for money. Recent academic literature has defined the difference between contract cheating and AI-based substitution as "cyber pseudepigraphy", or creating a false impression of authorship using electronic (digital) means, and 'AI-Giarism', or using AI tools to help create assignments that would

otherwise take longer to complete without assistance from technology [19], [20], which recognize that GenAI-mediated misconduct occupies a novel category requiring distinct analytical and policy responses.

2.2. Rationalization, peer norms, and institutional cues

There are four primary behavioral manifestations which are evident toward using GenAI in accomplishing academic work, as shown repeatedly in the literature. First, students' perceptions about the usefulness of AI will impact their level of incentive to utilize the technology for completing a task [21], [22]. The second element of social construction is peer norms and the behavior of educators; when many peers use the tools there is substantial social pressure, and opting out creates higher perceived costs of not using AI tools [15], [16]. Third, students deploy cognitive neutralization strategies, including denial of injury 'only improves my language'), denial of responsibility ('the guidelines were unclear'), and condemnation of condemners ('instructors are inconsistent about AI rules'), to resolve the moral dissonance that arises when their behavior crosses perceived ethical boundaries [23], [24]. Recent studies show that stronger endorsement of these neutralization strategies reliably predicts higher rates of general cheating as well as AI-specific misconduct [23]. Fourth, the clarity and consistency of institutional expectations around AI use strongly moderate whether rationalizations actually lead to rule-breaking behavior [25], [26]. In the absence of clear, unified guidance from the institution, students often read the resulting ambiguity as implicit approval [11], [26].

2.3. The Saudi higher education context

As part of the Kingdom of Saudi Arabia's Vision 2030, the higher education sector has been expanded rapidly. The digital transformation, workforce readiness, and the rapid expansion of educational modernization as part of the higher education sector, is part of the Kingdom's effort to create new ways of learning through a more diverse student body and an emphasis on using AI and considering the full possibilities of digital transformation [8], [9]. Students who are enrolled in English-medium instructional programs still face the problem of limited proficiency in English. GenAI has two functions for these students. First, generative AI can be used as a tool for providing a means of overcoming language barriers (e.g., through translation or explanation). Second, students can use generative AI as an effective method for managing their large volume of assignments and/or coursework [11], [12]. Recent research about Saudi English as a foreign language (EFL) learners suggests that feedback produced by ChatGPT leads to written improvement equivalent to that provided by human instructors [27]. The practical appeal of AI replacement will likely increase in the future as these models continue to develop further. At the same time, instructors frequently report having only a limited amount of experience with AI detection tools and inadequate education on how to adapt assessments for use in environments where generative AI is easily accessible [28], [29]. The overlap of these pressures on students, peers, norms, and institutional gaps underscores a need for empirical investigations of how motivational, normative, and structural factors affect AI-based replacement. This research project fills this gap in understanding by examining the perceptions of both students and instructors at a Saudi university which provides additional insights into NNES contexts around the globe, particularly those in which English medium instruction is a primary requirement for advancement within academia.

3. THEORETICAL FRAMEWORK

Through our research, we develop an understanding of AI substitution, through the establishment of a model, based on learning decisions, that considers motivation and normative influences of the student on these decisions; as well as the development of moral rationale for these decisions created by the institution's expectation. Rather than judging the use of GenAI as either ethical or unethical, this framework focuses on how students perceive their use of AI in the context of the performance expectations and boundaries that they see as acceptable. This dual grounding in normative and motivational theory provides a more nuanced explanation of boundary-crossing behaviors than would be the case with models of AI misuse based on either a strong relationship between attitudes and use or as a purely rule-breaking behavior.

3.1. Distinct construct definitions

To ensure conceptual clarity and minimize construct overlap, the four latent constructs in this study are distinguished along functional lines. Attitudes toward AI (L1) are defined as utility-driven performance appraisals: the cognitive evaluation of AI as useful, efficient, and performance-enhancing. These represent the 'what' of behavior, what students believe AI can do for them. Moral rationalization (L2) is defined as post-hoc cognitive neutralization of ethical conflict: the justification process through which students reconcile boundary-crossing behavior with a positive self-concept. This represents the 'how' of behavior, how students maintain identity as 'good students' while engaging in potentially prohibited conduct. Institutional and instructional clarity (L3) is defined as the regulatory environmental signal: the perceived

explicitness, consistency, and specificity of expectations regarding AI use. This represents the 'where' of behavior, the environmental context that either constrains or enables justification processes. AI-based substitution (L4) is defined as the procedural replacement of cognitive drafting phases through generative AI, distinct from contract cheating (transactional product outsourcing) and from legitimate scaffolding (supportive cognitive enhancement).

3.2. Motivational and normative drivers

Drawing on the TPB [30], the study posits that students' engagement with AI-based substitution is proximally predicted by attitudes (evaluations of utility), subjective norms (perceptions of peer use and faculty tolerance), and perceived behavioral control (perceived detection likelihood). In learning contexts characterized by time pressure, linguistic demands, or high-stakes grading, utility-driven motivation increases the attractiveness of AI use [22]. The TPB components become particularly relevant in the multicultural university context of Saudi Arabia, where linguistic and academic demands have caused students to perceive GenAI as a valuable tool for accomplishing their academic work.

3.3. Moral rationalization as boundary-crossing mechanism

In order to understand the cognitive influences behind engaging in ethically ambiguous behaviors, the framework incorporates neutralization theory [31]. In this study, the following specific neutralization strategies employed were denial of responsibility ('the directions were not clear'), denial of injury ('I just want to improve my language'), condemnation of the condemner ('the instructors are inconsistent'), and appeal to a higher loyalty ('I have to support my scholarship'), are all related to unique pressures associated with being a Saudi NNES. In addition, because of the language barrier, students are able to neutralize their use of AI-generated prose as 'language polishing', rather than as cognitive outsourcing; therefore, their denial of injury is related specifically to the highly multilingual nature of their academic environment. Recent empirical work confirms that endorsement of neutralization techniques significantly predicts AI-facilitated academic misconduct [23]. These findings extend the classical neutralization theory framework to contemporary digital learning environments, demonstrating that the moral-cognitive mechanisms Sykes and Matza [31] identified in juvenile delinquency operate with comparable explanatory power in GenAI-mediated boundary crossing.

3.4. Institutional clarity as contextual moderator

Institutional clarity is conceptualized as a contextual moderator that conditions the translation of rationalization into substitution behavior. When expectations are clear, consistent, and illustrated, the scope for moral rationalization shrinks and the apparent cost of boundary crossing grows. When expectations are vague or partial, students engage rationalization strategies on the assumption that institutional silence means tacit agreement. This mental model resonates with the integrity climate school of thought that clarity and perceived fairness drive compliance and trust [25], [26].

3.5. Research questions and hypotheses

The following research questions guide this investigation:

- To what extent do attitudes toward generative AI predict AI-based substitution? (RQ1)
- To what extent do moral rationalization strategies predict AI-based substitution? (RQ2)
- Does perceived institutional clarity moderate the rationalization–substitution relationship? (RQ3)
- How do students and instructors describe the learning, ethical, and contextual factors shaping perceptions of acceptable AI use? (RQ4)

Furthermore, there several hypotheses of this study: i) positive attitudes toward GenAI use are positively associated with AI-based substitution behavior (H1); ii) moral rationalization is positively associated with AI-based substitution behavior (H2); iii) institutional clarity is negatively associated with AI-based substitution behavior (H3); and iv) institutional clarity moderates the relationship between rationalization and substitution, such that the positive association is weaker when clarity is high (H4).

4. RESEARCH METHOD

4.1. Research design

This study adopted a sequential explanatory mixed-methods design [32]: we began with a quantitative phase (students' motivation measured by survey data and structural equation modeling or SEM) and followed with qualitative phase to further shed light on contextually how to interpret the quantitative findings, students' and teachers' accounts shed further light on the quantitative phases. We sequenced phases so that they triangulate meaningfully: student motivations (quantified with survey and SEM data) with pedagogical realities and tensions (lived by instructors and students; gleaned from our interview data).

However, this integration was more than the production of a corresponding dataset per phase that could be informative [13], [14]. That is, the quantitative phase identified important predictors of AI substitution and the qualitative phase delved into the whys of not just the effect of the predictors on substitution but why the sizes of the effects and moderating role of clarity of institutional roles on these relationships.

4.2. Conceptual model

The conceptual model comprises four latent constructs: attitudes toward AI (L1), moral rationalization (L2), institutional clarity (L3), and AI-based substitution (L4). The model posits that substitution is directly predicted by L1 and L2, with L3 acting as both a direct predictor and a moderator of the L2→L4 path. A graphical representation is provided in Figure 1.

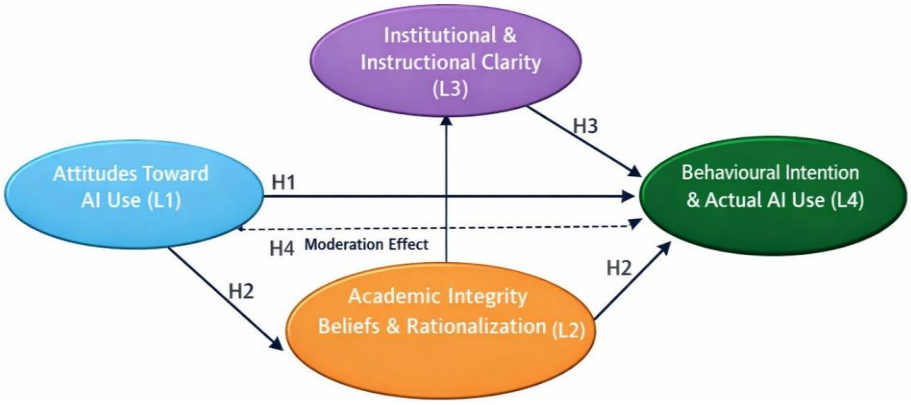

Figure 1. Conceptual model of AI-assisted academic behavior

4.3. Measures and operationalization

All constructs were assessed using reflective indicators with items measured on a 5-point Likert scale (1=strongly disagree; 5=strongly agree). Survey items were also adapted from well-established theoretical models, as well as from recent AI-in-education research. Content validity was evaluated by a panel of three expert scholars in applied linguistics and educational ethics. Following expert review, the rationalization items were specified to reflect the four classical neutralization strategies (denial of injury, denial of responsibility, condemnation of condemners, and appeal to higher loyalty) rather than general attitudes or concerns, and the substitution items were specified to capture observable substitution behavior rather than self-efficacy or detection anxiety, in order to ensure each indicator represents its intended construct. Table 1 contains the full item wording, standardized factor loadings, composite reliability (CR), and average variance extracted (AVE) across constructs.

Table 1. Latent construct operationalization and measurement properties

Construct	Item	Survey item wording	Loading	CR	AVE
Attitudes (L1)	L1_1	AI tools help me brainstorm and generate initial ideas for assignments	0.78	0.84	0.64
	L1_2	AI tools improve the quality and efficiency of my academic writing	0.82		
	L1_3	Using AI tools enhances my overall learning performance	0.79		
Rationalization (L2)	L2_1	Using AI to paraphrase my writing is acceptable because it does not change my own ideas (denial of injury)	0.71	0.81	0.59
	L2_2	When AI rules are unclear, students cannot be blamed for using AI to complete tasks (denial of responsibility)	0.77		
	L2_3	Using AI is fair when instructors are inconsistent about what is allowed (condemnation of condemners)	0.83		
Clarity (L3)	L3_1	My institution has clear policies regarding acceptable AI use	0.74	0.78	0.51
	L3_2	My instructors provide explicit guidelines on how AI may be used	0.69		
	L3_3	I feel comfortable discussing my AI use openly with instructors	0.73		
Substitution (L4)	L4_1	I use AI to generate substantial portions of my assignment content	0.80	0.86	0.63
	L4_2	I have submitted assignments in which AI generated significant portions of the text	0.77		
	L4_3	I rely on AI to draft sections of my work that I then submit as my own	0.84		

Note: All loadings significant at $p<0.001$.

4.4. Research context and participants

The study was conducted at a large public university in Saudi Arabia offering preparatory and undergraduate programmers through the medium of English. This program is situated in a multilingual unit where students experience high-stakes testing and adopt the same technologies adopted for classroom use in Vision 2030 [8]. Stratified purposive sampling method was employed to ensure variety in terms of gender, educational level or year, and discipline. Quantitative sample included 249 undergraduate students. Qualitative sample of 7 students (from survey participants who indicated willingness to participate) and 10 instructors teaching writing intensive classes.

4.4.1. Quantitative sample size justification

The adequacy of n=249 for partial least squares structural equation modeling (PLS-SEM) with moderation effects was assessed using the inverse square root method [33], the current standard for minimum sample size estimation in non-parametric structural modeling. Unlike fixed rules-of-thumb, this method derives sample requirements from the statistical properties of the model itself. To detect a minimum path coefficient of 0.158 at 80% statistical power (α=0.05), the formula yields:

$$N = \left(\frac{2.486}{0.158}\right)^2 \approx 247.5$$

With n=249, the calculated threshold is exceeded, supporting sufficient statistical power for the hypothesized interaction. This approach is consistent with current PLS-SEM guidance recommending model-specific sample size estimation rather than fixed cut-offs [34]. The sample therefore provides a methodologically sound basis for detecting the moderation effect specified in H4 with an acceptable probability of Type II error.

4.4.2. Qualitative sample justification

The qualitative phase employed 7 student interviews and 10 instructor interviews within a sequential explanatory design. In this framework, the purpose of interviews is not standalone statistical generalizability but to explain quantitative findings through participant perspectives [32]. The sample size was justified using the information power principle [35], [36], which holds that the adequacy of a qualitative sample depends on five dimensions: study aim (narrowly defined: AI substitution justifications), sample specificity (homogeneous: Saudi English as a second language (ESL) undergraduates), established theory (strong: TPB and neutralization theory guided analysis), quality of dialogue (high: 25–40 minute semi-structured interviews), and analysis strategy (cross-case: reflexive thematic analysis). Across these dimensions, information power was high, and thematic redundancy was observed by the sixth student interview, at which point the emergence of unique codes dropped below the 5% saturation threshold. For instructors, redundancy was observed by the eighth interview. Table 2 summarizes the qualitative sampling framework.

Table 2. Qualitative sample justification and saturation framework

Dimension	Students (n=7)	Instructors (n=10)	Strategy
Information power	High (specific context)	High (disciplinary diversity)	Contextual saturation
Design role	Sequential explanatory	Sequential explanatory	Explain quantitative β magnitudes
Analytical method	Reflexive thematic	Reflexive thematic	Identify latent meaning
Saturation marker	Redundancy by interview 6	Redundancy by interview 8	Comparative code analysis

4.5. Data collection procedures

The survey was administered via the university learning management system over a three-week period. Participation was voluntary and responses were anonymous. Interview participants were drawn from willing survey respondents and faculty teaching writing-intensive courses. All interviews were audio-recorded with consent, transcribed verbatim, and lasted between 25 and 40 minutes. Interviews were conducted face-to-face or via secure videoconferencing.

4.6. Data analysis

PLS-SEM [37], [38] was employed to test hypothesized relationships and moderation effects, given its appropriateness for exploratory behavioral models, non-normal distributions, and moderate sample sizes. Analysis proceeded in two stages: i) measurement model assessment (indicator loadings, CR, AVE, and heterotrait–monotrait (HTMT) ratios) and ii) structural model assessment (path coefficients, bootstrapped significance tests with 5,000 resamples, R^2 values, effect sizes f^2, and product-indicator moderation analysis). Interview transcripts were analyzed using reflexive thematic analysis [39]. Two analysts independently coded

for meaning units relevant to motivation, justification, and context, grouping codes into higher-order themes aligned with theoretical constructs while allowing context-specific themes to emerge. Integration occurred at the interpretation layer through narrative synthesis and joint display tables [13], [14].

4.7. Ethical considerations

Ethical approval was obtained from the English Language Institute Research Ethics Committee (approval number on file). The study was conducted in accordance with the Helsinki Declaration. Participation was voluntary, informed consent was obtained, and data were anonymized prior to analysis. Participation did not affect academic standing or coursework evaluation.

5. RESULTS AND DISCUSSION

5.1. Measurement model

The measurement model demonstrated adequate reliability and validity for all four latent constructs. CR values of 0.78 to 0.86 are greater than 0.70. AVE values of 0.51 to 0.64 confirmed convergent validity. All standardized indicator loadings were significant and greater than 0.60 (Table 1). Discriminant validity was confirmed using the HTMT criterion in all instances, with values falling comfortably below the conservative threshold of 0.85 [34]. A summary of the reliability and validity indicators across the four constructs is presented in Table 3. This evidence effectively addresses potential concerns about conceptual overlap between attitudes and rationalization: although both constructs are linked to AI use, they operate at distinct functional levels, attitudes reflecting utility appraisal, rationalization serving as moral justification, and the HTMT results demonstrate that they remain empirically separable.

Table 3. Measurement model reliability and validity summary

Construct	CR	AVE	HTMT (max)	Status
Attitudes (L1)	0.84	0.64	0.72	Valid
Rationalization (L2)	0.81	0.59	0.79	Valid
Clarity (L3)	0.78	0.51	0.83	Valid
Substitution (L4)	0.86	0.63	0.81	Valid

Note: HTMT threshold <0.85.

5.2. Structural model

The structural model explained 46% of observed variance in AI-based substitution behavior (R^2=0.46), representing a moderate level of explanatory power attributable to the combined effects of attitudes, rationalization, and institutional clarity. Path coefficient results confirmed statistically significant direct effects for all three predictors. Full structural results are presented in Table 4, and the estimated model is illustrated in Figure 2.

Table 4. Structural model results and hypothesis testing

Hypothesis	Path	β	SE	t	p	f^2	Result
H1	Attitudes → substitution	0.39	0.08	4.88	<0.001	0.21	Supported
H2	Rationalization → substitution	0.31	0.09	3.42	0.001	0.14	Supported
H3	Clarity → substitution	–0.27	0.10	2.70	0.007	0.09	Supported
H4	Rat. × Clarity → substitution	–0.22	0.11	2.05	0.041	0.07	Supported

Note: Bootstrap 5,000 resamples. β=standardized path coefficient; f^2=effect size.

5.3. Moderation analysis

The interaction effect of institutional clarity on the rationalization–substitution link was statistically significant (β=−0.22, p<0.05), supporting H4. Simple slopes analysis revealed that the association between rationalization and substitution was significantly weaker at high levels of perceived clarity compared to low clarity, where rationalization strongly predicted substitution. Figure 3 illustrates this interaction.

The moderation effect size (f^2=0.07) falls between Cohen's benchmarks for small (0.02) and medium (0.15) effects. While this magnitude warrants interpretive caution, three considerations support its theoretical significance. First, in complex human behavioral research involving moderation of attitudinal–behavioral pathways, interaction effects are systematically attenuated by measurement error and restricted variance, making f^2 values above 0.02 practically meaningful [34]. Second, the moderation captures a contextual constraint on a moral-cognitive process; even small shifts in the clarity–rationalization dynamic

can produce cumulative institutional effects when scaled across student cohorts over time. Third, the effect size is consistent with moderation findings in recent educational psychology and integrity-climate research examining how environmental cues regulate self-regulatory behavior [25]. Thus, the interaction is interpreted as a small-to-moderate effect with meaningful policy implications rather than as a trivial finding.

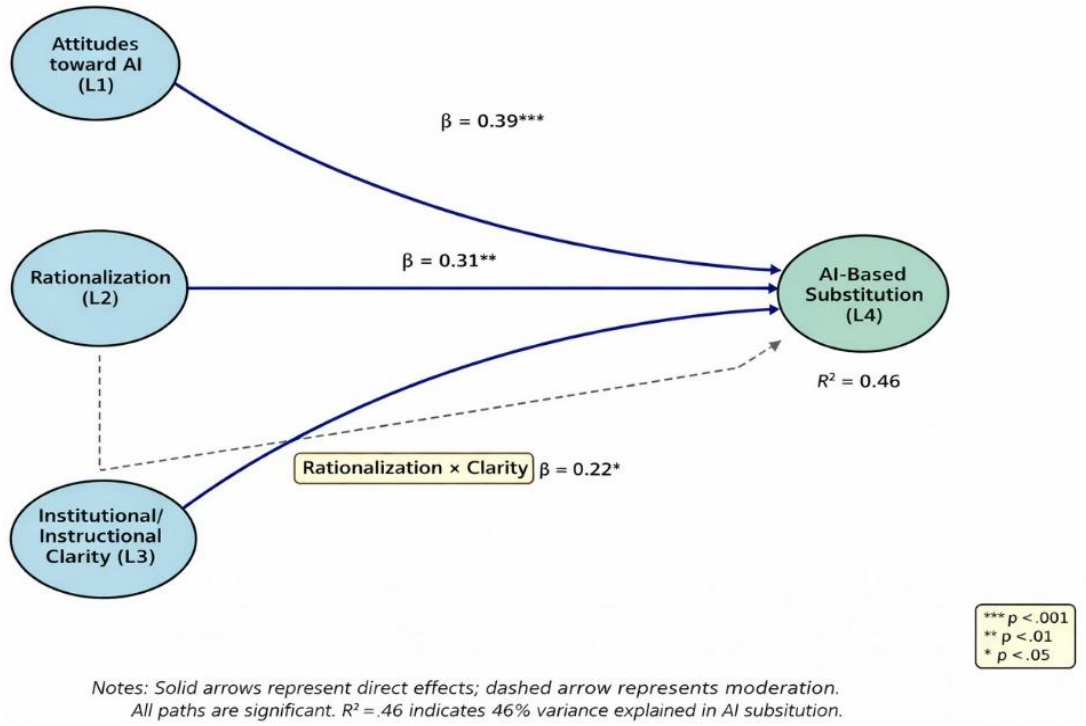

Figure 2. Structural equation model with standardized path coefficients

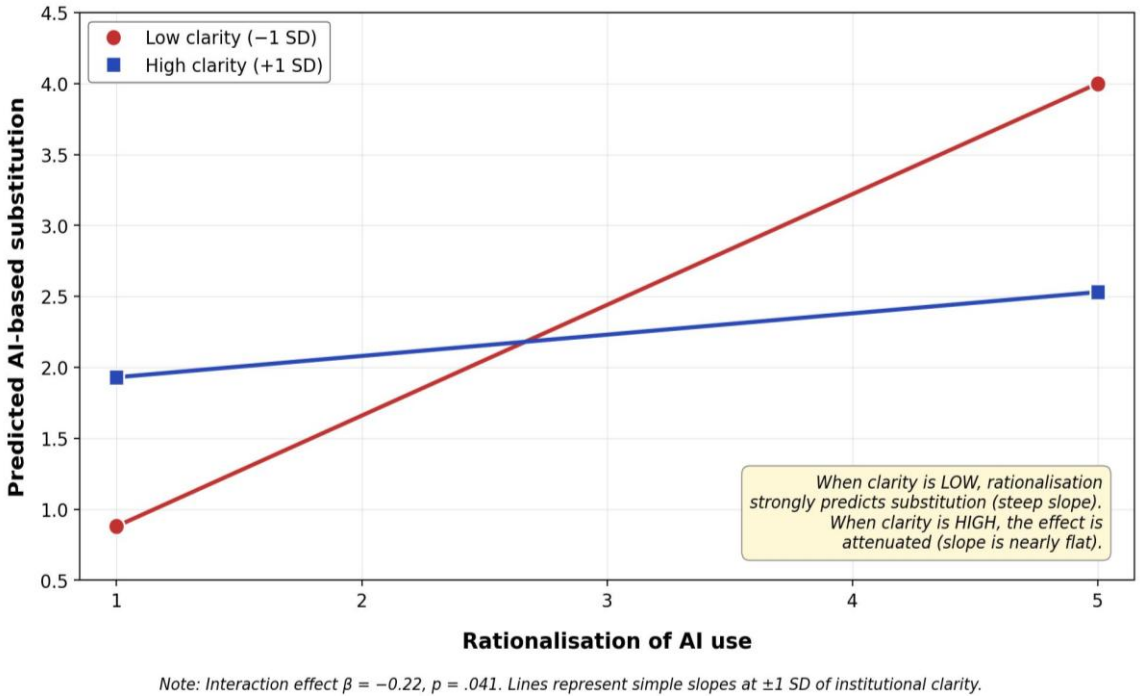

Figure 3. Moderation of rationalization–substitution relationship by institutional clarity

5.4. Qualitative results

The analysis identified three major themes that illustrate the ways motivational pressures, ethical considerations and contextual cues enhanced or inhibited the way that AI is used by students (and their instructors) in their academic work. These three themes were developed inductively through reflexive thematic analysis of both student and instructor interviews, mapped to the theoretical constructs of attitudes, rationalization, and institutional clarity. Together, they provide contextual depth that substantiates and extends the statistical findings from the structural model.

5.4.1. Theme 1: AI as cognitive and emotional support

According to students, AI has reduced the feeling of anxiety associated with starting heavily written assignments and provided them with support to get started on an academic task. One student said: "*There are times when I have full knowledge of my subject, but I can't get my ideas down on paper until I've looked at a sample*." Students were also clear that for multilingual students, AI assists with the ability to get through language barriers, such as, "*AI helps me think of how to put my ideas together and also helps me write the proper academic word*." A number of students voiced their concerns that they would become dependent upon AI as they see immediacy in terms of productivity will conflict with developing skills long-term. This dual function, AI as both a "starter kit" and a potential "crutch", contextualizes the strong attitudinal path (β=0.39) by revealing that utility-driven motivation is intrinsically tied to workload management and linguistic burden.

5.4.2. Theme 2: ethical ambiguity and boundary negotiation

Many students were uncertain about when they could use and rely on AI since they rationalize their understanding of AI in terms of how it works in practice rather than on principle. One participant reported that he was not cheating because he had his own thoughts and was just changing things around. Another participant reported that if someone could finish their assignment with AI, then the assignment itself could be the problem. Social norms played a significant role of student conceptions of AI help acceptability. "*Everybody does it. If you don't, you will just make it harder for yourself*." These patterns confirm that rationalization operates as a moral buffer enabling boundary crossing (β=0.31), with the specific neutralization strategies of denial of injury emerging as the primary mechanism in NNES contexts where students frame AI use as 'language polishing' rather than cognitive outsourcing.

5.4.3. Theme 3: institutional silence and fragmented guidance

Instructors expressed frustration with inconsistent messaging across departments; some permitted AI use while others prohibited it, leading to ad hoc individual guidelines. Faculty members expressed limited confidence in the AI detection tools used to identify student work and their predictions (AI detection tools) of how difficult it would be to validate assessment results. Students reported hiding the use of AI because of perceived inconsistency between their instructors regarding how or whether AI should be used; as one student stated, "*I have had one instructor that has allowed the use of AI and another that has not mentioned it, so I would prefer to keep this to myself.*" The systemic nature of instructor responses is shown in Table 5.

Table 5. Categorization of instructor interview findings

Category	Representative finding	Freq. (n/10)	Implication
Assessment redesign	Difficulty designing tasks resistant to AI substitution	7/10	Process-based assessment models
Detection limitations	Low confidence in AI-detection accuracy for NNES writing	8/10	Shift from surveillance to transparency
Policy ambiguity	No departmental guidance; individual instructor discretion	6/10	Standardized AI-use policies required
Student concealment	Awareness that students hide AI use	9/10	Open disclosure norms needed

5.5. Mixed-methods integration: joint display of meta-inferences

Through the application of integration methodology [13], [14], we used quantitative structural pathways, paired with qualitative themes to develop 'meta-inferences', i.e., insights only derived from integrating both data types. A joint view of integrated findings is shown in Table 6, along with the integration of interviews, providing contextual and nuanced interpretations of all statistical effects. In this way, we have produced a series of elaborated 'meta-inferences' [13] about the relationships between the different types of data.

Table 6. Joint display of integrated meta-inferences

Quantitative effect	Qualitative theme	Integrated meta-inference	Recommendation
Attitudes → Sub. (β=0.39***)	AI as efficiency and coping tool	Students assess how AI can support/assist them in decreasing barriers to academic labor and how, as complexity increases, a scaffold can become a shortcut with limited pedagogical support.	Scaffold drafting stages explicitly in assessments.
Rat. → Sub. (β=0.31**)	Ethical ambiguity and boundary negotiation	Denial of injury is the primary mechanism for NNES boundary crossing; students frame substitution as "language polishing" rather than cognitive outsourcing.	Provide explicit NNES language support and clear substitution examples.
Clarity → Sub. (β=–0.27**)	Institutional silence and fragmented guidance	Institutional silence is interpreted as tacit permission; absence of guidance does not produce neutrality but actively enables rationalization.	Implement standardized AI disclosure policies.
Clarity×Rat. (β=–0.22*)	Policy inconsistency amplifies peer norms	Ambiguity acts as a catalyst for moral neutralization; when one instructor permits and another is silent, students default to the permissive norm.	Align departmental and course-level guidelines.

Note: ***p<.001; **p<.01; *p<.05.

5.6. Global transferability of findings

While rooted in the Saudi higher education context, our findings demonstrate considerable analytical transferability to other NNES environments worldwide. The mechanisms identified in this study, utility-oriented attitudes amplified by language barriers, rationalization facilitated by lack of institutional clarity, and managing through explicit clarity, are not culturally specific but structurally characteristic of any higher education context where English-medium instruction, English-language publication or assessment as a criterion for degree or career advancement, and widely available generative AI tools. The confluence of these trends can also be seen in countries that belong to the Gulf Cooperation Council (GCC), East Asia, and regions of Sub-Saharan Africa, where there is increasing reliance on high stakes tests for students who are primarily NNES, as well as the rapid digital transformation of education through technology. The integrated theoretical framework, combining the TPB with neutralization theory and institutional clarity as a key moderator, thus provides a robust, portable lens for investigating contested scaffold–shortcut boundaries in AI use wherever these conditions apply [19], [20], [26].

6. CONCLUSION

This research examined the influences on students' generative AI substitution behavior at Saudi universities, researching the roles of motivational appraisals, moral rationalization and institutional clarity, through a mixed-methods sequential design that utilized PLS-SEM and surveys together with qualitative interviews. The structural model demonstrated that positive attitudes and rationalization strategies combined to explain 46% of the variance in substitution behavior, with the pathway from rationalization to substitution significantly reduced by institutional clarity, affirming the theoretical assertion that transparency in policy serves as a moral de-risking mechanism. In addition, qualitative data yielded rich descriptions of how denial-of-injury neutralization specifically occurs in NNES contexts, with students framing AI-generated text as "language-polishing", not as cognitive outsourcing, while instructors reported systemic barriers to coherent policy and assessment redesign.

This study contributes an analytical lens for assessing the effectiveness of institutional AI guidance, one that explicitly models the interaction between moral neutralization and institutional clarity. The framework builds on existing technology acceptance models while foregrounding ethical and contextual moderators of AI-related behavior. On the practical side, the findings support layered governance that distinguishes acceptable scaffolding from unacceptable substitution, alongside assessments that foreground visible learning processes over end products. They also point to AI fluency that positions students as critical editors rather than passive consumers of AI outputs, and to culturally responsive integrity education attuned to the linguistic demands and performance pressures faced by multilingual learners.

The key limitations of the current study are reliance on self-reported measures of behavior and restriction to a single institution within one cultural setting. These constraints affect generalizability and may attenuate effect sizes through common-method variance. Longitudinal work is therefore needed to examine changes over time, and cross-cultural, multi-institutional studies could investigate developmental patterns and the role of differing institutional interventions.

Ultimately, preserving educational integrity in the age of generative AI is likely to require regulation and pedagogy as much as technology. Institutions that model their policies, approaches to teaching, and assessment practices on principles of transparency, active cognitive engagement, and ethical agency will be

best situated to harness the advantages of AI while shielding 'real' learning. Helping students learn to develop responsible usage of AI will not only be an integrity goal but also one of modern academic literacy and competence.

ACKNOWLEDGEMENTS

The authors wish to thank the students and instructors who participated in this study.

FUNDING INFORMATION

No external funding was received for this research.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Mazin Mansory	✓	✓			✓				✓			✓	✓	
Zilal Meccawy				✓		✓		✓		✓				

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

INFORMED CONSENT

Written informed consent was obtained from all student and instructor participants prior to data collection. Participants were informed of the study's aims, their right to withdraw at any time without penalty, the anonymity of responses, and the secure storage of data. No personally identifying information is reported.

ETHICAL APPROVAL

This study was approved by the English Language Institute Research Ethics Committee, King Abdulaziz University, and conducted in accordance with the Declaration of Helsinki.

DATA AVAILABILITY

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

REFERENCES

[1] H. Balalle and S. Pannilage, "Reassessing academic integrity in the age of AI: a systematic literature review on AI and academic integrity," *Social Sciences and Humanities Open*, vol. 11, p. 101299, 2025, doi: 10.1016/j.ssaho.2025.101299.

[2] K. Bittle and O. El-Gayar, "Generative AI and academic integrity in higher education: a systematic review and research agenda," *Information*, vol. 16, no. 4, p. 296, Apr. 2025, doi: 10.3390/info16040296.

[3] K. D. Wang, Z. Wu, L. Tufts, C. Wieman, S. Salehi, and N. Haber, "Scaffold or crutch? Examining college students' use and views of generative AI tools for STEM education," in *2025 IEEE Global Engineering Education Conference (EDUCON)*, Apr. 2025, pp. 1–10, doi: 10.1109/EDUCON62633.2025.11016406.

[4] M. Bearman, J. Tai, P. Dawson, D. Boud, and R. Ajjawi, "Developing evaluative judgement for a time of generative artificial intelligence," *Assessment & Evaluation in Higher Education*, vol. 49, no. 6, pp. 893–905, Aug. 2024, doi: 10.1080/02602938.2024.2335321.

[5] W. Liang, M. Yuksekgonul, Y. Mao, E. Wu, and J. Zou, "GPT detectors are biased against non-native English writers," *Patterns*, vol. 4, no. 7, p. 100779, Jul. 2023, doi: 10.1016/j.patter.2023.100779.
[6] M. Darban, "The future of virtual team learning: navigating the intersection of AI and education," *Journal of Research on Technology in Education*, vol. 57, no. 3, pp. 659–675, May 2025, doi: 10.1080/15391523.2023.2288912.
[7] Y. R. Marín *et al.*, "Artificial intelligence as a teaching tool in university education," *Frontiers in Education*, vol. 10, p. 1578451, Apr. 2025, doi: 10.3389/feduc.2025.1578451.
[8] I. K. Al-Dosari, "The role of continuous education in achieving intellectual security in the Kingdom of Saudi Arabia," *Educational Sciences: Theory and Practice*, vol. 25, no. 2, pp. 124–147, 2025, doi: 10.12738/jestp.2025.2.09.
[9] E. Faisal, "Unlock the potential for Saudi Arabian higher education: a systematic review of the benefits of ChatGPT," *Frontiers in Education*, vol. 9, p. 1325601, Apr. 2024, doi: 10.3389/feduc.2024.1325601.
[10] I. Glendinning, "Responses to student plagiarism in higher education across Europe," *International Journal for Educational Integrity*, vol. 10, no. 1, pp. 4–20, May 2014, doi: 10.21913/IJEI.v10i1.930.
[11] A. Alammari, "Evaluating generative AI integration in Saudi Arabian education: a mixed-methods study," *PeerJ Computer Science*, vol. 10, p. e1879, Feb. 2024, doi: 10.7717/peerj-cs.1879.
[12] D. M. Bamasoud, R. Mohammad, and S. Bilal, "Adopting generative AI in higher education: a dual-perspective study of students and lecturers in Saudi universities," *Big Data and Cognitive Computing*, vol. 9, no. 10, p. 264, Oct. 2025, doi: 10.3390/bdcc9100264.
[13] A. Younas, S. Fàbregues, S. Munce, and J. W. Creswell, "Framework for types of metainferences in mixed methods research," *BMC Medical Research Methodology*, vol. 25, no. 1, p. 18, Jan. 2025, doi: 10.1186/s12874-025-02475-8.
[14] Y. Zhou, Y. Zhou, and K. Machtmes, "Mixed methods integration strategies used in education: a systematic review," *Methodological Innovations*, vol. 17, no. 1, pp. 41–49, Mar. 2024, doi: 10.1177/20597991231217937.
[15] D. L. McCabe and L. K. Trevino, "Academic dishonesty: honor codes and other contextual influences," *The Journal of Higher Education*, vol. 64, no. 5, p. 522, Sep. 1993, doi: 10.2307/2959991.
[16] D. L. McCabe, L. K. Trevino, and K. D. Butterfield, "Cheating in academic institutions: a decade of research," *Ethics & Behavior*, vol. 11, no. 3, pp. 219–232, Jul. 2001, doi: 10.1207/S15327019EB1103_2.
[17] A. S. Espartinez, "Exploring student and teacher perceptions of ChatGPT use in higher education: a q-methodology study," *Computers and Education: Artificial Intelligence*, vol. 7, p. 100264, Dec. 2024, doi: 10.1016/j.caeai.2024.100264.
[18] K. J. Topping, E. Gehringer, H. Khosravi, S. Gudipati, K. Jadhav, and S. Susarla, "Enhancing peer assessment with artificial intelligence," *International Journal of Educational Technology in Higher Education*, vol. 22, no. 1, p. 3, Jan. 2025, doi: 10.1186/s41239-024-00501-1.
[19] C. K. Y. Chan, "Students' perceptions of 'AI-Giarism': investigating changes in understandings of academic misconduct," *Education and Information Technologies*, vol. 30, no. 6, pp. 8087–8108, Apr. 2025, doi: 10.1007/s10639-024-13151-7.
[20] A. K. Kofinas, C. H. H. Tsay, and D. Pike, "The impact of generative AI on academic integrity of authentic assessments within a higher education context," *British Journal of Educational Technology*, vol. 56, no. 6, pp. 2522–2549, Nov. 2025, doi: 10.1111/bjet.13585.
[21] C. K. Y. Chan and W. Hu, "Students' voices on generative AI: perceptions, benefits, and challenges in higher education," *International Journal of Educational Technology in Higher Education*, vol. 20, no. 1, p. 43, 2023, doi: 10.1186/s41239-023-00411-8.
[22] Ö. F. Ursavaş, Y. Yalçın, H. İslamoğlu, E. Bakır-Yalçın, and M. Cukurova, "Rethinking the importance of social norms in generative AI adoption: investigating the acceptance and use of generative AI among higher education students," *International Journal of Educational Technology in Higher Education*, vol. 22, no. 1, p. 38, Jun. 2025, doi: 10.1186/s41239-025-00535-z.
[23] J. Hawdon, M. Costello, and A. V. Reichelmann, "Cheating with ChatGPT and techniques of neutralization," *Deviant Behavior*, vol. 47, no. 4, pp. 664–685, Apr. 2026, doi: 10.1080/01639625.2025.2456067.
[24] C. Ainsworth, W. A. Chernoff, Y. J. Chae, and M. Bisciglia, "Examining academic dishonesty in online group chats through the lenses of strain and neutralization theories," *Deviant Behavior*, vol. 45, no. 3, pp. 361–376, Mar. 2024, doi: 10.1080/01639625.2023.2248337.
[25] J. Janinovic, S. Pekovic, R. Djokovic, and D. Vuckovic, "Assessing the effectiveness of academic integrity institutional policies: how can honor code and severe punishments deter students' cheating—moderating approach?" *SAGE Open*, vol. 14, no. 4, pp. 1–16, Oct. 2024, doi: 10.1177/21582440241307430.
[26] S. Wang, F. Wang, Z. Zhu, J. Wang, T. Tran, and Z. Du, "Artificial intelligence in education: a systematic literature review," *Expert Systems with Applications*, vol. 252, p. 124167, Oct. 2024, doi: 10.1016/j.eswa.2024.124167.
[27] A. H. Alsofyani and A. M. Barzanji, "The effects of ChatGPT-generated feedback on Saudi EFL learners' writing skills and perception at the tertiary level: a mixed-methods study," *Journal of Educational Computing Research*, vol. 63, no. 2, pp. 431–463, Apr. 2025, doi: 10.1177/07356331241307297.
[28] A. Amigud and C. Ellis, "The great unknown: faculty responses to generative AI and its impact on assessment," *Assessment & Evaluation Higher Education*, vol. 48, no. 8, pp. 1189–1205, 2023, doi: 10.1080/02602938.2023.2223842.
[29] D. R. E. Cotton, P. A. Cotton, and J. R. Shipway, "Chatting and cheating: ensuring academic integrity in the era of ChatGPT," *Innovations in Education and Teaching International*, vol. 61, no. 2, pp. 228–239, Mar. 2024, doi: 10.1080/14703297.2023.2190148.
[30] I. Ajzen, "The theory of planned behavior," *Organizational Behavior and Human Decision Processes*, vol. 50, no. 2, pp. 179–211, Dec. 1991, doi: 10.1016/0749-5978(91)90020-T.
[31] G. M. Sykes and D. Matza, "Techniques of neutralization: a theory of delinquency," *American Sociological Review*, vol. 22, no. 6, pp. 664–670, Dec. 1957, doi: 10.2307/2089195.
[32] J. W. Creswell and V. L. Clark, *Designing and conducting mixed methods research*, 3rd ed. Thousand Oaks, CA: SAGE Publications, Inc., 2017.
[33] N. Kock and P. Hadaya, "Minimum sample size estimation in PLS-SEM: the inverse square root and gamma-exponential methods," *Information Systems Journal*, vol. 28, no. 1, pp. 227–261, Jan. 2018, doi: 10.1111/isj.12131.
[34] J. F. Hair, M. Sarstedt, C. M. Ringle, P. N. Sharma, and B. D. Liengaard, "Going beyond the untold facts in PLS–SEM and moving forward," *European Journal of Marketing*, vol. 58, no. 13, pp. 81–106, Dec. 2024, doi: 10.1108/EJM-08-2023-0645.
[35] K. Malterud, V. D. Siersma, and A. D. Guassora, "Sample size in qualitative interview studies," *Qualitative Health Research*, vol. 26, no. 13, pp. 1753–1760, Nov. 2016, doi: 10.1177/1049732315617444.
[36] A. Wutich, M. Beresford, and H. R. Bernard, "Sample sizes for 10 types of qualitative data analysis: an integrative review, empirical guidance, and next steps," *International Journal of Qualitative Methods*, vol. 23, pp. 1–14, Jan. 2024, doi: 10.1177/16094069241296206.

[37] J. F. Hair, C. M. Ringle, and M. Sarstedt, "Partial least squares structural equation modeling: rigorous applications, better results and higher acceptance," *Long Range Planning*, vol. 46, no. 1–2, pp. 1–12, Feb. 2013, doi: 10.1016/j.lrp.2013.01.001.
[38] J. Henseler, C. M. Ringle, and R. R. Sinkovics, "The use of partial least squares path modeling in international marketing," in *Advances in International Marketing*, vol. 20, R. R. Sinkovics and P. N. Ghauri, Eds., Bingley, UK: Emerald Group Publishing Limited, 2009, pp. 277–319, doi: 10.1108/S1474-7979(2009)0000020014.
[39] V. Braun and V. Clarke, "Using thematic analysis in psychology," *Qualitative Research in Psychology*, vol. 3, no. 2, pp. 77–101, Jan. 2006, doi: 10.1191/1478088706qp063oa.

BIOGRAPHIES OF AUTHORS

Mazin Mansory is an associate professor of TESOL and Dean of the English Language Institute at King Abdulaziz University, Jeddah, Saudi Arabia. He earned his Ph.D. in TESOL from the University of Exeter, United Kingdom in 2017, following postgraduate studies in applied linguistics. His research interests span language assessment, second-language teacher education, professional development, teacher beliefs and cognition, and the responsible integration of emerging technologies in English-medium higher-education contexts. He has published in peer-reviewed venues on assessment literacy, teacher beliefs, and curriculum reform, and serves on institutional committees concerned with academic standards and quality assurance. In the present study, he led the conceptualization, methodological design, formal analysis, and overall project administration, ensuring rigor from instrument development through the integration of the mixed-methods findings. He can be contacted at email: mmmansory@kau.edu.sa.

Zilal Meccawy is an assistant professor of Education at the English Language Institute, King Abdulaziz University, Jeddah, Saudi Arabia. She earned her Ph.D. from the University of Nottingham, United Kingdom, where her doctoral research engaged with language acquisition and the psychology of the language learner. Her ongoing research interests include language assessment, bilingualism in higher education, technology-enhanced language learning, learner motivation, and the role of generative AI in shaping multilingual academic literacy. She has presented at international conferences on educational psychology and technology adoption and teaches academic English in writing-intensive undergraduate courses. In the present study, she contributed to investigation, data curation, validation, and writing through review and editing, with a particular focus on the qualitative interview phase and on integrating learner perspectives with the structural findings. She can be contacted at email: zmeccawy@kau.edu.sa.