International Journal of TESOL Studies (2026)
260602, 1-21 https://doi.org/10.58304/ijts.260602

Article

Exploring Multilingual Learners' AI-mediated Informal Digital English-speaking Practice: A Q-methodology Study

Li Rong
Murad Abdu Saeed*
Universiti Malaya, Malaysia

Received: 1 March 2026/Received in revised form: 5 May 2026/Accepted: 23 May 2026/ Available online: 18 June 2026

Abstract
Despite the increasing use of artificial intelligence (AI) tools in informal English-speaking practice, there is still limited understanding of multilingual learners' subjective perceptions and engagement patterns in this context. To fill this gap, the current study aimed to identify core subjective factors and explore how learners with different first language (L1) backgrounds and English proficiency levels differ in their perception and engagement patterns. The study replicated the Q-methodology of Guo and Xia (2025), who explored learners' perception of AI-mediated informal digital multilingual learning, in investigating perception of AI-assisted informal speaking practice (AI-IDESP) among multilingual undergraduate students in Malaysia. Using the Extended Unified Theory of Technology Acceptance and Use (UTAUT2), and a Q-Methodology, the present study surveyed and interviewed 30 multilingual undergraduate students in a public Malaysian university who had engaged in AI-IDESP and use at least one other language in daily communication. The findings show that performance-oriented use of AI, emotion-driven motivation and efficiency-driven but anxiety dependence influenced learners' behavioural intentions and actual engagement with AI-IDESP. Students' different English proficiencies were more likely to align with these different factors. Overall, by deepening our understanding of how learners' AI-IDESP is shaped by performance orientation, emotional sensitivity, and efficiency-related anxiety, the study offers useful implications for learner-centred AI tool design in multilingual contexts.

Keywords
English language learning, Q-methodology, AI-IDESP, AI speaking practice, multilingual learners

1 Introduction

Informal digital language learning (IDLL) has become an important part of contemporary language education, reflecting learners' greater control over their learning goals and encounters with the target language in real-life situations (Lee & Lee, 2021; Richards, 2015). Yet, English speaking practice

*Corresponding author. Email: muradsaeed@um.edu.my

remains challenging for many ESL/EFL learners to maintain outside the classroom. Most of these learners often struggle to find low-pressure, interactive opportunities for speaking practice (Horwitz et al., 1986; Lee & Lee, 2021). Speaking in English is also influenced by emotional factors, including foreign language anxiety (FLA) (Horwitz et al., 1986). Students may avoid speaking in English because they feel embarrassed about making mistakes, worry about losing face, or expect negative judgement from others (Horwitz et al., 1986). As a result, speaking in English continues to be one of the most challenging skills to develop through IDLL, despite the wide availability of digital learning resources (Lee & Lee, 2021; Richards, 2015).

These challenges make informal speaking practice a good context in which artificial intelligence (AI) may be especially relevant. Unlike many conventional digital resources, AI-based speaking tools can provide immediate feedback, opportunities for interaction, and on-demand conversational practice in a relatively low-pressure environment (Xing & Saeed, 2025; Yang & Li, 2024). Such affordances may help learners sustain oral practice beyond the classroom, especially when they have limited access to suitable speaking partners or feel anxious about being judged by humans for pronunciation, grammar, and vocabulary mistakes. Against this background, recent studies have highlighted new opportunities for informal speaking practice through AI-based speech technologies such as ChatGPT, Elisa, and pronunciation-focused applications which can provide immediate responses, simulate conversations, and act as readily available speaking partners (Liu et al., 2025; Yang & Li, 2024). Speaking practice with AI can help address learners' limited access to suitable conversation partners and fear of social judgement, thus providing more accessible opportunities for oral practice (Horwitz et al., 1986; Yang & Li, 2024). A growing body of research has shown that AI-mediated informal digital learning of English (AI-IDLE) supports more flexible and personalized learning (Liu et al., 2025; Soyoof et al., 2021), although concerns remain about inaccurate information, unnatural language output, and the risk of learners becoming overly dependent on AI for language use (Yu et al., 2025; Zhang & Liu, 2022). Recent AI-IDLE scholarship has also argued for a skill-specific perspective, noting that different language skills involve distinct challenges and psychological mechanisms, and therefore may be supported differently by AI tools (Zou et al., 2025).

Yet, most existing research on AI-IDLE has been carried out in relatively monolingual contexts, such as the Chinese EFL context where learners share a dominant first language (L1) and use English in relatively uniform linguistic settings (e.g., Liu et al., 2025; Liu et al., 2024). In contrast, multilingual contexts such as Malaysia present different conditions for informal English-speaking practice where learners have different first languages (Bahasa Melayu, Mandarin, Tamil, and other minority languages) and they often switch between languages (Gill, 2013). Therefore, AI-mediated English-speaking practice (AI-IDESP) is likely to be closely connected with multilingual learning behaviours, such as asking AI for explanations in the L1, comparing expressions across languages, and dealing with pronunciation issues related to accent and cross-linguistic influence (Liu et al., 2024; Yu et al., 2025). In such contexts, learners' L1 backgrounds may matter not because they mechanically determine attitudes, but because they shape how learners draw on multilingual resources when interacting with AI tools. Recent research on informal digital learning of languages other than English (IDL-LOTE) also highlights the need to recognize that learners' engagement in IDLL is shaped by motivation, enjoyment, self-confidence, and the wider language environment (Liu et al., 2024).

However, little research has examined how multilingual learners in contexts like Malaysia perceive and engage in their AI-IDESP. This gap is important because previous IDLE research has more often emphasized its benefits than learners' reluctance to participate, even though hidden barriers such as unsupportive environments and low motivation may constrain engagement (Nguyen, 2026). Therefore, drawing on the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2) (Venkatesh et al., 2003; Venkatesh et al., 2012), this study adopts Q-methodology to identify shared configurations of viewpoints among multilingual Malaysian undergraduates by replicating recent work on multilingual AI-IDLE (Guo & Xia, 2025). In this study, UTAUT2 is used as a conceptual framework to identify key dimensions shaping learners' perceptions of AI-IDESP, such as usefulness, ease of use, facilitating conditions, and habit. Combined with Q-methodology, this approach helps examine not only which beliefs matter, but also how these beliefs cluster into distinct subjective viewpoints. It addresses the following research questions:

1. What are the core subjective factors influencing the behavioural intentions and actual participation levels of Malaysian multilingual undergraduates in AI-IDESP?
2. How do learners with different L1 backgrounds and English proficiency levels differ in their perceptions and engagement patterns toward AI-IDESP?

2 Literature Review

2.1. Informal digital language learning and speaking practice

IDLL generally refers to learners' attempts to learn a language by themselves beyond the formal classroom setting, using digital tools (Lee, 2018, 2019; Richards, 2015). Learners frequently use digital tools for receptive exposure, for example, watching English YouTube videos, listening to songs, or browsing English social media, but engage less in productive activities such as voice or video interaction in English (Lee, 2019; Soyoof et al., 2021). Recent work has also indicated that informal digital English learning should not be treated as a unitary construct because different activity types may involve distinct cognitive processes and may contribute differently to language development (Zou et al., 2026). However, sustaining speaking practice in informal environments is more challenging since it requires synchronous interaction, and real-time responses (Lee, 2018, 2019).

The affective dimension (i.e., emotions such as anxiety and enjoyment) is central to understanding this difficulty. Based on the FLA construct by Horwitz and Cope (1986), research has revealed that FLA is associated with lower willingness to communicate, avoidance of oral participation, and reduced perceived competence (Dewaele & MacIntyre, 2014; Dewaele et al., 2018; MacIntyre et al., 1998). Similar emotional patterns can emerge in informal digital contexts, where learners may worry about losing face, or being judged for pronunciation or grammatical errors (Lee, 2019; Liu, 2019; Soyoof et al., 2021). Thus, enjoyment and motivation are crucial for sustaining engagement in IDLL. According to Lee and Lee (2021), both engagement in IDLE and learners' ideal L2 self significantly predicted foreign language enjoyment (FLE). Moreover, diversity and quality of IDLE activities are positively related to motivation, self-confidence, and reduced speaking anxiety (Lee, 2018). However, speaking remains a major challenge because it is socially visible and emotionally demanding, especially for university students (Dressman & Sadler, 2020; Lee, 2019).

In addition to fear of negative evaluation, oral production places substantial real-time processing demands on learners. For example, speaking requires simultaneous attention to idea generation, lexical retrieval, grammatical encoding, pronunciation, and self-monitoring, all under time pressure. When AI is introduced into this process, learners may experience an additional burden of evaluating whether AI-generated language is accurate, natural, and contextually appropriate (Liu, Wang, & Zou, 2025). Therefore, learners' engagement in AI-mediated speaking practice may be shaped not only by enjoyment or usefulness, but also by speaking anxiety and cognitive load associated with monitoring both their own language production and the AI's output (Zhang et al., 2025).

2.2. AI-mediated informal digital English learning and speaking

The rapid development of AI conversational agents has expanded AI-IDLE (Liu et al., 2024; Soyoof et al., 2021; Yang & Li, 2024). Learners increasingly use AI assistants for dialogue practice, role-play, feedback on grammar and vocabulary, and explanations of language use in everyday out-of-class contexts (Yang & Li, 2024; Yu et al., 2025). Unlike earlier digital resources with static input, these AI systems can act as quasi-conversational partners and on-demand language advisers in learners' daily routines (Lee & Lee, 2021). Learners are also able to adopt AI tools not only for instrumental reasons, but also for comfort, enjoyment, and alignment with personal needs (e.g., Yu et al., 2025). Moreover, AI tools can be beneficial for learners who lack regular chances to connect with proficient speakers of English in their local environment (Fang et al., 2025). They may also help address limited access and high affective pressure (Fang et al., 2025). Conversational AI can provide 24/7 access to a responsive partner in English (Fang et al., 2025; Liu et al., 2024). Moreover, practicing with a non-human partner can reduce fear of negative evaluation (Horwitz et al., 1986; MacIntyre et al., 1998).

Nonetheless, AI-mediated speaking practice creates new risks and uncertainties such as inaccuracies and over-reliance among learners who lack critical digital literacies (Soyoof et al., 2021; Yang & Li, 2024). AI-generated language may sound overly formal or "robotic," or may not reflect local pragmatic norms, and may be culturally distant from the interactions that learners actually face (Liu et al., 2024; Yu et al., 2025). Such limitations in AI feedback mechanisms and potential bias in training data contribute to more cautious attitudes, even among students who recognize the tools' efficiency (Fang et al., 2025; Soyoof et al., 2021). This indicates that learners may diversely perceive the benefits of AI-mediated speaking practice. Some learners focus primarily on efficiency and learning outcomes, others prioritize emotional comfort and anxiety reduction, and still others remain cautious due to concerns about accuracy, authenticity, or ethics (Yang & Li, 2024; Yu et al., 2025). To capture these differing positions, there is a need for methods that can reveal learners' subjective viewpoints and priorities in how they understand AI-based speaking practice.

2.3. Multilingual learners in IDLL and AI-IDLE

Empirical studies on IDLL and AI-IDLE have been conducted in contexts where learners share a dominant L1 and English as a foreign language, for example, in Chinese and Korean university settings (Lee, 2019; Lee & Lee, 2021; Liu et al., 2024). This focus has advanced the field but left an important gap for multilingual environments such as Malaysia where university students frequently use different first languages, and English learning is not isolated from other languages, but is embedded in a complex multilingual repertoire (Gill, 2013).

Recent scholarship has drawn attention to a "global English bias" in IDLL research and called for more work on IDL-LOTE (Liu et al., 2024; Soyoof et al., 2021). For instance, Liu et al. (2024) showed that promotion-focused motivation, enjoyment, and self-efficacy predicted participation in IDL-LOTE. In multilingual contexts, learners often use AI tools in strategically multilingual ways, such as asking AI to explain English vocabulary or grammar through their L1, requesting translations between multiple languages, or checking meaning differences via bilingual prompts (Liu et al., 2024; Yang & Li, 2024). Learners also evaluate whether AI systems recognize their local accents and varieties, such as Malaysian English, and whether the output aligns with local pragmatic and cultural norms (Gill, 2013; Soyoof et al., 2021). These practices are crucial because AI's impact depends less on its technical affordances and more on how learners integrate it into their existing language repertoires and everyday communicative situations (Liu et al., 2024; Yang & Li, 2024). While multilingual learners may find AI-IDESP supportive, they may also worry that AI privileges "standard" inner-circle English norms, under-represents local varieties, or misinterprets culturally embedded meanings, which could undermine perceived authenticity and even create new forms of discomfort (Fang et al., 2025; Gill, 2013; Liu et al., 2024; Soyoof et al., 2021; Yu et al., 2025).

These issues provide a clear rationale for examining subjective viewpoints among multilingual undergraduates in Malaysia. In this sense, learners' L1 background may matter not because it directly determines attitudes, but because it also shapes how learners use multilingual resources when interacting with AI. In AI-IDESP, learners may rely on their L1 to request explanations, compare meanings across languages, interpret feedback, and evaluate accent- or pragmatics-related issues (Liu et al., 2024; Yang & Li, 2024; Yu et al., 2025). In multilingual contexts such as Malaysia, language use is shaped by diverse linguistic repertoires and local English varieties, which makes L1 background a relevant contextual dimension for understanding variation in learners' perceptions and engagement (Gill, 2013; Soyoof et al., 2021).

2.4. Technology acceptance research and UTAUT2 in AI-mediated language learning

Technology acceptance models explain learners' adoption or rejection of new tools (Venkatesh et al., 2003; Venkatesh et al., 2012). Among these, UTAUT and UTAUT2 are widely applied in educational technology research. In language education, UTAUT-based studies typically test the factors which best predict learners' intention to use a given technology. Research reported that behavioural intention is influenced by perceived usefulness, with more context-dependent roles for social influence and facilitating conditions (Agyei & Razi, 2022).This logic has been extended to AI-related language learning contexts. Recent work on AI-assisted learning and large language models (LLMs) has reported that acceptance-related factors predict students' intentions to use LLMs in the future (An et al., 2023; Wang, 2024). Studies on AI anxiety and behavioural intention similarly show that perceived usefulness, attitudes, and trust shape whether students are willing to integrate AI into their learning practices (An et al., 2023). These findings are important because they determine factors which predict adoption or rejection of new tools.

However, survey-based acceptance research has limits when the goal is to understand learner subjectivity. UTAUT-type models usually break down attitudes into separate measurable variables and estimate average effects across a sample, which can conceal the way individual learners combine beliefs in complex, sometimes internally inconsistent configurations (Venkatesh et al., 2012). In practice, two learners may report similar usefulness scores for an AI tool but for quite different reasons, with some prioritizing instrumental outcomes and others placing greater emphasis on emotional comfort or concerns about potential risks. Traditional regression models can include multiple predictors, but they do not necessarily reveal how beliefs are combined into distinct patterns of viewpoint, especially when emotions and identity are involved (Watts & Stenner, 2005).This limitation is particularly salient for speaking practice, where affective factors such as anxiety, enjoyment, and self-image can shape behaviour as much as perceived usefulness (Horwitz et al., 1986; Lee & Lee, 2021). As a result, while UTAUT2 is useful for specifying the topics which should be covered, additional methods are necessary to understand why learners with similar acceptance scores may still engage with AI in different ways (Watts & Stenner, 2005; Yu et al., 2025).

In the present study, UTAUT2 is employed not as a regression-based predictive model, but as a conceptual framework that guided statement construction and factor interpretation. In particular, the Q-set was designed to reflect core UTAUT2-related dimensions, such as performance expectancy, effort expectancy, facilitating conditions, social influence, hedonic motivation, price value, and habit. At the same time, since AI-IDESP is affectively situated in multilingual contexts, the study remains open to dimensions that are not fully captured by UTAUT2 alone, including emotional authenticity, anxiety, privacy concerns, and fear of overreliance. In this way, Q-methodology makes it possible to examine how these beliefs and concerns are configured into shared learner viewpoints rather than treated as isolated predictors.

2.5 Q-methodology for learner viewpoints

Q-methodology aims to study subjectivity by identifying shared viewpoints' patterns among individuals (Brown, 1993). A typical Q study is initiated by constructing a set of statements (the Q-set) covering a wide range of positions on the topic, and then asking participants to sort these statements on a structured grid, by following a continuum from "most disagree" to "most agree," which compels them to express priorities (Watts & Stenner, 2005). The completed Q-sorts are analysed using by-person factor analysis to group participants with similar sorting patterns; each factor is interpreted as a distinct shared viewpoint rather than as a conventional variable-based trait (Brown, 1993; Watts & Stenner, 2005). This fits AI-mediated speaking practice well that learners often experience both advantages and concerns (Soyoof et al., 2021). More importantly, Q-methodology is particularly suitable for the present study since the aim is not to measure the net effect of isolated predictors, but to identify how learners combine beliefs, emotions, and concerns into shared subjective viewpoints. Unlike variable-centred approaches, which analyse relationships among variables across a sample, Q-methodology uses by-person factor analysis to reveal patterned configurations of subjectivity across participants (Brown, 1993; Watts & Stenner, 2005). This methodology, therefore, offers an important epistemological advantage for research on AI-mediated speaking practice, where learners may simultaneously perceive AI as useful, emotionally comforting, risky, or dependence-inducing.

In language education and applied linguistics, Q-methodology is applied to investigate identities, and mindsets, especially pre-service EFL teachers (Irie et al., 2018) and learners (Guo & Xia, 2025; Yu et al., 2025). Although UTAUT2-based surveys help identify key predictors (An et al., 2023; Venkatesh et al., 2012; Wang, 2024), they do not by themselves explain how these factors cohere into distinct learner viewpoints in multilingual, high-stakes speaking contexts (Brown, 1993; Lee & Lee, 2021; Watts & Stenner, 2005). Therefore, UTAUT2-based surveys- Q-methodology combination will offer a strong methodological and theoretical basis for the present study on multilingual Malaysian undergraduates' perceptions of AI-IDESP.

In the present study, the extracted factors refer to the empirical groupings produced through Q-factor analysis, while each factor is interpreted as representing a shared viewpoint among participants.

3 Method

3.1 Research design

This study used Q-methodology to investigate multilingual learners' subjective perceptions of AI-IDESP. This approach was selected because the study aimed to identify shared configurations of learner subjectivity rather than estimate the strength of relationships among pre-defined variables. Q-methodology focuses on how individuals rank a common set of statements relative to one another, thus revealing patterns of subjectivity and grouping participants with shared perspectives into factors (Guo & Xia, 2025; Irie et al., 2018; Watts & Stenner, 2005). Moreover, UTAUT2 was applied as a conceptual guide to ensure that the Q-set covered a broad range of dimensions. Specifically, the framework informed statement development across key dimensions such as performance expectancy, effort expectancy, social influence, facilitating conditions, hedonic motivation, price value, and habit, while additional statements captured affective and multilingual concerns identified in recent AI-IDLE research. This design aligns with previous work that has combined UTAUT-based constructs with Q-methodology to examine learners' technology-related beliefs and AI-mediated IDLL (An et al., 2023; Guo & Xia, 2025; Liu et al., 2025).

3.2 Participants

Participants were multilingual undergraduate students enrolled at one public Malaysian university. Since sampling in Q-methodology research aims to include participants who are information-rich and directly relevant to the research questions (Brown, 1993; Watts & Stenner, 2005), the current study used purposive sampling. To be eligible, students had to (a) use English and at least one other language (e.g., Malay, Mandarin, Tamil) in their daily lives, (b) have used AI tools for informal English-speaking practice within the previous six months, and (c) be willing to do the Q-sorting task and a short interview. Recruitment was conducted through a Google Form distributed face-to-face on campus. The researcher introduced the study, obtained informed consent, and collected basic demographic information, including students' year of their undergraduate studies, L1 backgrounds, self-reported English proficiency, and experience with AI tools for English speaking practice. Thirty undergraduates completed the full procedure, which is consistent with methodological recommendations for Q-studies, where samples of 20–40 participants are typical (Brown, 1993; Watts & Stenner, 2005). Thus, the study sample consisted of 30 undergraduates, with a higher proportion of female participants (N= 19; 63.3%) compared to male participants (N= 11; 36.7%). Most of them were in their first year of study (N=16; 53.3%), followed by those who were in their second year (N=8; 26.7%), while a smaller number of them reported that they were in their third (N=2; 6.7%) and fourth year of study (N=4; 13.3%). Regarding respondents' L1, 15 of them reported Bahasa Melayu (50%), and a substantial portion reported Mandarin (N=12; 40%) and Tamil (N=3; 10%) as their L1. For their self-reported English proficiency, respondents were split between advanced/high (N=16; 53.3%) and intermediate/medium (N=14; 46.7%) levels, with no participants indicating a beginner/low proficiency level (0%).

3.3 Data collection

Data collection followed standard Q-methodological procedures: (a) the Q-sorting process, and (c) post-sorting interviews:

3.3.1. Construction of the Q-set

A total of 30 valid Q-sorts were collected and analysed. The Q-set consisted of 34 statements representing learners' perceptions and experiences of AI-IDESP. Its development followed a theory-driven and literature-based process consistent with established Q-methodological principles (Brown, 1993; Watts & Stenner, 2005). Approximately, 60 candidate statements reflecting affective and contextual aspects were constructed based on UTAUT2 and previous relevant empirical research (An et al., 2023; Lee & Lee, 2021; Richards, 2015; Yang & Li, 2024). To ensure theoretical coverage, the statement pool included items related to perceived usefulness for speaking development and task performance (performance expectancy), ease of learning and using AI tools for conversation (effort expectancy), influence from significant others and peer trends (social influence), access to devices and internet resources (facilitating conditions), enjoyment and role-play interest (hedonic motivation), willingness to pay for premium tools (price value), and routine use of AI for speaking practice (habit). In addition, statements were included on anxiety, privacy, authenticity, and multilingual prompting practices, as these concerns are especially relevant in AI-mediated informal English-speaking practice (Yu et al., 2025). Table (1) presents an illustrative mapping of the Q-statements to the UTAUT2 and additional context-specific dimensions, providing the theoretical grounding for the statement-level findings reported later in Table 3.

The initial statement pool was refined by the researcher. Following Q-methodological guidelines on clarity, breadth of coverage, and practical feasibility (Brown, 1993; Guo & Xia, 2025; Watts & Stenner, 2005), semantically overlapping or redundant statements were merged or removed, and wording was simplified to enhance scale readability. Attention was paid to retaining a balance of supportive, neutral, and critical positions toward AI-IDESP. Through this process, 34 statements formed the final Q-set, which was considered sufficient to represent major performance-oriented, affective, and risk-aware dimensions of learners' perception while remaining manageable for completion within a single Q-sorting session.

Table 1

UTAUT2-informed Dimensions Represented in the Q-set

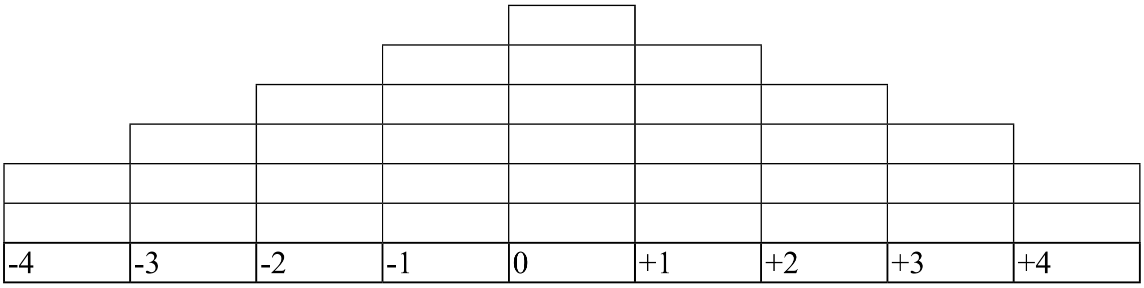
UTAUT2 dimension	Focus in this study	Example statement numbers
Performance expectancy	Usefulness for speaking improvement, test/interview preparation	1, 3, 17, 33, 34
Effort expectancy	Ease of learning and using AI in spontaneous speaking	4, 5, 30
Social influence	Influence of professors, mentors, student trends	6, 7
Facilitating conditions	Access to smartphone, internet, resources	8
Hedonic motivation	Fun, enjoyment, role-play entertainment	10, 11
Price value	Willingness to pay, free versions sufficient	12, 13
Habit	Daily use, routine engagement	14
Beyond UTAUT2	Anxiety, privacy, authenticity, multilingual practices	15–32, especially 20–31

3.3.2. Q-sorting procedure and interviews

The Q-sorting process was conducted offline between December 2025 and January 2026 on the campus. Participants completed the Q-sorting task during individual face-to-face sessions. They received written instruction on the purpose and procedure of the Q-sort. Each of the 34 statements was printed on a physical card measuring 32 mm × 22 mm and used for hands-on sorting. All participants followed the same two-step Q-sorting procedure. First, they read all statements and sorted them into three preliminary categories: "agree," "neutral/uncertain," and "disagree." Second, they placed the 34 statements onto a forced distribution grid which ranges from −4 (strongly disagree) to +4 (strongly agree), with a fixed number of slots at each level (Figure 1). The guiding prompt for the sorting task was: "To what extent does each statement reflect my own experience of using AI tools for informal English-speaking practice beyond the classroom? "This forced-choice distribution required participants to make relative judgments among statements and is a core feature of Q-methodology (Brown, 1993; Watts & Stenner, 2005). Each Q-sorting session lasted approximately 10–20 minutes. Upon completion, participants reviewed their final distribution to ensure that it accurately reflected their personal views.

Immediately after the Q-sorting task, a brief 5–10-minute post-sorting interview was conducted with each participant to support factor interpretation (Guo & Xia, 2025; Watts & Stenner, 2005). Although relatively short, these interviews were highly focused because participants had already expressed their priorities through the completed Q-sort. The interview protocol concentrated on the subjective meanings of salient placements, especially statements positioned at +4, −4, and 0, as well as any statements that participants found difficult to rank. Typical prompts included: "Why did you place this statement here?", "Which statements best reflect your experience?", and "Were any statements difficult or conflicting to sort?" This targeted design helped elicit concise but meaningful explanations linked directly to each participant's ranked configuration rather than broad general opinions. All interviews were audio-recorded with consent and transcribed verbatim via the Otter website.

Figure 1
Forced Distribution Grid (−4 to +4) Used for the Q-sorting Task



-4 is most disagree +4 is most agree

3.4 Data analysis

Data analysis followed established Q-methodological procedures, integrating both quantitative and qualitative components. All Q-sort data were entered into Ken-Q Analysis Desktop Edition (KADE) for statistical analysis. A correlation matrix of participants' Q-sorts was first generated, after which principal component analysis (PCA) was conducted to extract factors, followed by varimax rotation to obtain a clear and interpretable factor structure (Watts & Stenner, 2008) (Figure 2). PCA was selected as a practical and transparent extraction method for identifying shared patterns across participants' Q-sorts, while varimax rotation was used to achieve a simpler and more interpretable factor structure. This combination is widely used in contemporary Q-methodology research, especially when the goal is to obtain clearly defined factors for subsequent qualitative interpretation (Watts & Stenner, 2008; Guo & Xia, 2025).

After examining the scree plot, factor solutions ranging from one to eight factors were tested. The final number of factors retained was chosen using multiple criteria, including eigenvalues greater than 1.0, the requirement that each factor be defined by at least two Q-sorts with statistically significant loadings, the proportion of variance explained, and the overall interpretability of the factor solution, in line with recent Q-studies in applied linguistics and AI-IDLE (Guo & Xia, 2025). Factor loadings were also evaluated at the $p < 0.01$ significance level. Q-sorts that loaded significantly on a single factor were treated as defining sorts, whereas confounded and non-significant sorts were excluded from factor interpretation.

Figure 2
Scree Plot of Eigenvalues for Factor Extraction

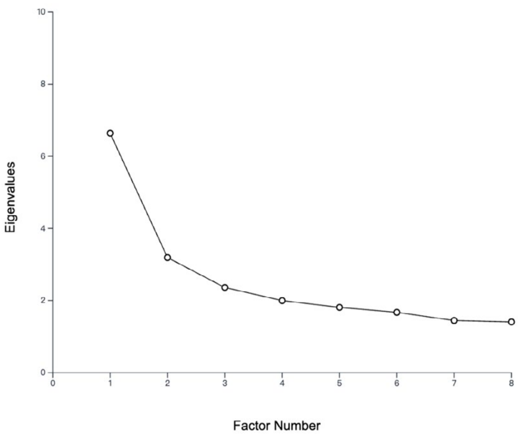


Table 2
Q-sort Factor Loadings of Participants

Participant	Factor
P01	F2
P02	F2
P03	F1
P04	(confounded)
P05	F1
P06	F1
P07	F1
P08	(non-significant)
P09	F1
P10	(non-significant)
P11	F3
P12	F2
P13	F3
P14	F1
P15	F1
P16	(non-significant)
P17	F1
P18	F1
P19	F1
P20	(non-significant)
P21	F2
P22	F1
P23	(confounded)
P24	(non-significant)
P25	F1
P26	F2
P27	(non-significant)
P28	(confounded)
P29	F1
P30	F1

After examining the scree plot and testing factor solutions ranging from one to eight factors, a three-factor solution was retained (Figure 2). The rotated factors together accounted for 41% of the total variance, with Factor 1 explaining 22%, Factor 2 explaining 11%, and Factor 3 explaining 8%. This level of explained variance meets and slightly exceeds the 35%–40% threshold commonly considered acceptable in Q methodology (Watts & Stenner, 2008), indicating that the extracted factor structure is statistically sound and interpretable.

With regard to factor loading criteria, the significance level was set at $p < 0.01$, corresponding to a critical loading value of ±0.456. Q-sorts with loadings greater than +0.456 or less than −0.456 on a single factor were regarded as defining sorts and were used for subsequent factor interpretation. Q-sorts that loaded significantly on more than one factor (confounded sorts) as well as those that did not load significantly on any factor (non-significant sorts) were excluded from factor interpretation. Interview transcripts were used alongside the factor arrays and distinguishing statements to clarify the subjective meanings underlying extreme and salient statement placements, thereby supporting the interpretation and naming of the extracted factors.

Before presenting detailed interpretations of each factor, the distribution of participants across the factors is summarized. In the Q-sort, Factor 1 was defined by 14 participants, Factor 2 by 5 participants, and Factor 3 by 2 participants. In addition, six participants produced Q-sorts that were classified as non-significant sorts (Q08, Q10, Q16, Q20, Q24, and Q27), while three participants produced Q-sorts that loaded significantly on more than one factor and were thus classified as confounded sorts (Q4, Q23, and Q28). Both non-significant and confounded sorts were excluded from the final factor interpretation. For reference, Table (2) presents the factor loadings of all participants, as well as those classified as non-significant or confounded.

4 Results

The factor array scores for all statements are presented in Table 3 and the results are presented in the subsequent sub-sections.

Table 3

Statements and Factor Arrays for AI-IDLL

Statement Number	Statements	Factor 1	Factor 2	Factor 3
1	I find AI tools useful for improving my English speaking skills in my daily life.	3**	0	−2
2	AI tools help me correct my pronunciation errors effectively.	1*	0	−4**
3	If I use AI, I will increase my chances of getting better grades in English speaking tests.	0	−2	-1
4	Learning how to use AI tools for speaking practice is easy for me.	0	0	0
5	I find AI tools easy to use for spontaneous English conversation.	2	-3**	3
6	People who influence my behavior (e.g., professors, mentors) think I should use AI for speaking practice.	−4**	−1	1
7	Using AI for English practice is considered a trend among university students.	−1	1	-2
8	I have the necessary resources (e.g., smartphone, stable internet) to use AI for speaking practice.	3	-1	3
9	I engage in online communities (e.g., Xiaohongshu, Reddit) to learn tips about using AI for English.	−3**	1	0
10	Using AI to practice English speaking is fun.	2**	−1	−1
11	I find it entertaining to explore different role-play scenarios with AI.	−2	2**	4
12	Even if I have to pay for a subscription (e.g., ELSA Pro, ChatGPT Plus), I would do it for the speaking practice benefits.	−3**	0	2**

Statement Number	Statements	Factor 1	Factor 2	Factor 3
13	The free versions of AI tools are sufficient for my daily English speaking practice.	1	-2	0
14	Using AI for speaking practice has become a daily habit for me.	−2**	−4**	2**
15	I feel less anxious speaking to an AI than to a real human because it doesn't judge me.	0	−2**	1
16	I feel safe practicing pronunciation with AI because I don't have to worry about "losing face."	−2	−4	-1
17	Using AI helps me build confidence before I speak English in real-life situations.	2**	−1	−3
18	I appreciate that AI never gets tired or impatient, no matter how many times I repeat a sentence.	2	1	−1*
19	Sometimes, talking to an AI feels lonely because there is no real human connection.	0	3**	0
20	I often use my mother tongue (e.g., Malay/Mandarin) to ask the AI to explain English concepts.	−3*	0	−1
21	I feel AI understands my "Malaysian English" or specific accent better than some native speakers.	0	0	2**
22	My multilingual background makes it easier for me to switch languages when interacting with AI.	1	1	2
23	I like that AI can translate difficult English expressions into my native language instantly.	3*	1**	3
24	Sometimes AI gives me answers that sound correct but are actually culturally inappropriate.	−2	3**	−3
25	I worry that AI might teach me unnatural or "robotic" English expressions.	−1**	4**	−4**
26	I worry that relying too much on AI will make me lazy in thinking for myself.	4	4	−3**
27	I feel mentally exhausted after a long session of talking to an AI.	−4**	2	0
28	I am concerned about data privacy when recording my voice for AI apps.	1	3**	1
29	The voice of the AI sometimes sounds too artificial and lacks emotion.	0	2**	0
30	I find that setting up good prompts for the AI takes too much time and effort.	−2	2**	−2
31	I feel frustrated when the AI misunderstands my voice or accent repeatedly.	−1	-2	1*
32	The "24/7 availability" is the main reason I choose AI over human partners.	−1**	−3**	4**
33	I specifically use AI to prepare for oral presentations or interviews.	4**	-1	−2
34	I use AI primarily to expand my vocabulary for daily conversation.	1	−3**	1

* Distinguishing statement at $p < 0.05$

* * Distinguishing statement at $p < 0.01$

4.1 Subjective Factors in AI-IDLE

Factor analysis of the Q-sorts revealed three factors in the context of English learning. In this study, these factors are treated as the empirical outcomes of the Q-method analysis, each representing a shared learner viewpoint. The labels assigned to the factors and the later discussion of learners' logics of practice are interpretive descriptions derived from these shared viewpoints. Factor analysis of subjective statements using the Q-methodology revealed three main factors in the context of English learning: performance-oriented use of AI-IDESP (Factor 1), emotion-driven engagement with AI (Factor 2), and efficiency-driven but anxiety dependence on AI (Factor 3). Table (4) presents the Q-sorts with significant loadings for each factor.

Table 4
Q-sorts with Significant Loadings for Each Factor

Factor	Q-sort
Factor 1: Performance-oriented use of AI	P03, P05, P06, P07, P09, P14, P15, P17, P18, P19, P22, P25, P29, P30
Factor 2: Emotion-driven engagement with AI	P01, P02, P12, P21, P26
Factor 3: Efficiency-driven but anxiety dependence on AI	P11, P13

4.1.1 Factor 1: performance-oriented use of AI

Factor 1 explains 22% of the total variance and reflects an AI use orientation primarily driven by learning performance and efficiency enhancement. Learners associated with this factor perceive AI as a highly instrumental support resource, mainly used to improve English speaking performance, accomplish specific learning tasks, and prepare for high-stakes language production contexts such as oral presentations and interviews. Their motivation for AI use is clearly dominated by performance expectancy.

Within this factor, learners strongly acknowledge the practical value of AI in improving English speaking skills (Statement 1: +3), and particularly emphasize its role in task-oriented speaking activities such as preparing presentations or interviews (Statement 33: +4). They generally believe that they possess the necessary objective conditions for using AI in speaking practice, including smart devices and stable internet access (Statement 8: +3), and highly appreciate AI's "tireless and always patient" nature (Statement 18: +2). In addition, these learners perceive AI as helpful in building speaking confidence prior to real-life communication (Statement 17: +2), and moderately recognize its ease of use in spontaneous speaking interaction (Statement 5: +2) as well as the enjoyment derived from using AI (Statement 10: +2). The interview data further corroborate this logic. For example, P03 reported that she mainly used ChatGPT or Gemini to "find alternative expressions to replace commonly used words," in order to expand her vocabulary and avoid repetitive expressions. Similarly, P15 noted that before presentations or interviews, she deliberately relied on AI to organize and optimize her language output. These accounts indicate that learners in Factor 1 do not regard AI as a substitute for their own thinking, but rather position it as an auxiliary tool to enhance their linguistic quality and expressive precision.

At the same time, learners in Factor 1 demonstrate clear distancing from social influence and cost-related considerations. They strongly reject the idea that significant others (e.g., teachers or mentors) expect them to use AI for speaking practice (Statement 6: −4), and rarely rely on online communities to learn AI usage tips (Statement 9: −3), suggesting that their AI use is driven primarily by personal learning needs rather than external normative pressure. Moreover, they are generally unwilling to pay for additional AI subscriptions for speaking practice (Statement 12: −3) and do not consider AI-based speaking practice a daily habit (Statement 14: −2), reflecting a rational and selective rather than dependency-oriented use pattern.

Notably, Factor 1 exhibits a coexistence of high efficiency recognition and strong self-regulatory awareness. Although these learners highly value AI's convenience in instant translation and vocabulary comprehension (Statement 23: +3), perceiving it as a time-saving tool, they strongly worry that overreliance on AI may weaken their independent thinking ability (Statement 26: +4). In interviews, several participants explicitly expressed discomfort with the tendency to "go straight to AI instead of thinking first" when answering questions or completing tasks. Although they firmly deny that prolonged interaction with AI leads to mental exhaustion (Statement 27: −4), they remain cautious about potential risks such as data privacy (Statement 28: +1), "mechanical" language output (Statement 25: −1), and the lack of emotional expressiveness in the AI voice (Statement 29: 0).

Overall, Factor 1 represents an AI use orientation centred on performance enhancement, emphasizing efficiency while maintaining strong critical awareness. These learners regard AI as an efficient and controllable language learning tool rather than a form of emotional companionship or long-term dependence. While actively leveraging AI's advantages, they remain acutely aware of the risks of cognitive laziness and overreliance, demonstrating a high level of self-regulation and instrumental rationality.

4.1.2 Factor 2: emotion-driven engagement with AI

Factor 2 explains 11% of the variance and is characterized by learners' heightened attention to their emotional experiences and affective responses. This motivational orientation does not stem from traditional external social influences, but instead it is primarily driven by emotional safety, anxiety reduction, and affective feedback experienced during human–AI interaction. Learners in this factor are highly sensitive to potential emotional and linguistic risks when using AI.

Specifically, they expressed strong concerns that AI may produce unnatural or "mechanical" English expressions (Statement 25: +4), and agreed that AI sometimes provides responses that are pragmatically or culturally inappropriate (Statement 24: +3). As stated by a representative participant, P21 (factor loading = 0.4773), "AI lacks empathy" and cannot display emotions or natural reactions in the way real human interaction does, which leads him to consistently prefer human interlocutors over AI.

At the same time, learners in Factor 2 exhibit pronounced emotional discomfort and anxiety. They agree that prolonged interaction with AI can induce feelings of loneliness (Statement 19: +3), and that AI voices often sound overly artificial and emotionally flat (Statement 29: +2). For instance, participant P02 (factor loading = 0.5528) remarked that compared with AI, humans are far superior in generating ideas and expressing emotions, noting that AI "finds it hard to truly understand emotions," which makes him reluctant to treat AI as a primary speaking partner.

In addition, learners in this factor demonstrate strong ethical and emotional risk awareness in their AI use. They expressed considerable concern about voice data privacy (Statement 28: +3). For example, P12 (factor loading = 0.576) stated that although AI is helpful in organizing language and expressing emotions, she is clearly aware that AI "stores a lot of personal information," which undermines her sense of security.

Importantly, despite their emotional sensitivity, learners in Factor 2 do not entirely reject the functional value of AI. They generally acknowledge that AI can support language learning to some extent; however, they strongly disagree that "24/7 availability" is the primary reason for choosing AI (Statement 32: −3), and they do not consider AI-based speaking practice a daily habit (Statement 14: −4). For example, P01 (factor loading = 0.5946) explicitly stated that although he uses AI to expand vocabulary or replace expressions, he does not regard AI as an indispensable practice partner. Overall, Factor 2 reflects an emotion-centred AI use motivation characterized by heightened sensitivity to potential risks.

4.1.3 Factor 3: Efficiency-driven but anxiety dependence on AI

Factor 3 explains 8% of the total variance and it is characterized by learners' strong recognition of AI's efficiency and convenience, accompanied by pronounced anxiety, vigilance, and self-restraint at both emotional and cognitive levels. This factor reflects a persistent tension between efficiency reliance and risk concern.

Learners in Factor 3 acknowledged AI's efficiency advantages in providing immediate support, such as translation and other language support (Statement 23: +3). However, such use is primarily contextual and compensatory rather than indicative of sustained learning engagement. They also valued AI's ease of use for spontaneous English conversation (Statement 5: +3) and its 24/7 availability (Statement 32: +4). At the same time, they explicitly rejected the effectiveness of AI in correcting pronunciation errors (Statement 2: −4). Notably, they strongly disagreed that AI might teach unnatural or "robotic" English expressions (Statement 25: −4), suggesting that their reservations were more selective than broad-based. Overall, this factor reflects strong recognition of AI's accessibility and convenience, accompanied by caution about specific limitations rather than generalized distrust.

The interview data vividly illustrates this "efficiency–anxiety coexistence." For example, participant P13 (factor loading = 0.6002) noted that although AI is highly efficient in providing information and facilitating understanding, its output often sounds rigid and unnatural, which may lead to stereotyped expressions. This concern over linguistic quality prompts him to remain highly cautious about AI's role in language learning. He also emphasized that "24/7 availability" is not his main reason for using AI, preferring instead to treat AI as a necessary but cautiously used tool.

Similarly, P11 (factor loading = 0.4795) stressed that excessive reliance on AI may result in cognitive laziness and weaken learners' ability to think independently and engage in authentic human interaction. For him, authentic language development is ultimately based on human-to-human communication rather than prolonged interaction with AI. These views align closely with the factor's strong negative loadings on several statements, further highlighting the anxiety and self-restraint embedded in this use orientation.

Overall, Factor 3 represents an AI use pattern that is efficiency-driven yet strongly regulated by anxiety and risk awareness. Learners do not deny AI's practical value in specific tasks, but their use is consistently accompanied by reflection on linguistic naturalness, cognitive degradation, and overdependence risks. This factor reveals that efficiency advantages in AI-IDLE do not necessarily translate into positive emotional engagement, and may instead generate complex anxiety-dependent reliance. From a theoretical perspective, this factor reflects a mixed configuration of UTAUT2-related beliefs. On the one hand, learners recognized strong performance expectancy and facilitating conditions, particularly the immediate availability and practical support offered by AI. On the other hand, such positive evaluations seem to coexist with low trust in AI's linguistic naturalness and with pronounced anxiety about dependence, reduced independent thinking, and inaccurate pronunciation feedback. This finding implies that, in AI-mediated speaking practice, positive acceptance beliefs do not necessarily produce confident or habitual use. Instead, they may coexist with speaking-related anxiety and heightened monitoring demands, indicating that UTAUT2 alone is insufficient unless complemented by affective and cognitive perspectives.

4.2 English proficiency of participants across factors

Table (5) presents the distribution of participants' self-reported English proficiency levels across the three extracted factors. Among participants loading significantly on Factor 1 (n =14), the majority of them reported advanced/high English proficiency (n =9), while the remaining participants identified as having intermediate proficiency (n =5). A similar pattern was observed for Factor 2 (n =5), where most participants also reported advanced proficiency (n =4). In contrast, participants associated with Factor 3 (n =2) were exclusive to the intermediate proficiency group.

Overall, these results suggest that learners with higher English proficiency were more likely to align with Factor 1, which emphasizes the performance-oriented and instrumental use of AI for English learning. This tendency is also reflected in participants' qualitative accounts. For example, one advanced-level participant described using AI primarily to expand vocabulary and explore alternative lexical choices in daily communication (P03), thus highlighting a goal-oriented and efficiency-driven approach to AI-IDESP. Similarly, several advanced learners reported using AI strategically for test preparation, pronunciation feedback, and vocabulary refinement rather than for emotional support or companionship.

Table 5

Distribution of Participants' English Proficiency across Factors

English proficiency level	**Factor 1 (n=14)**	**Factor 2 (n=5)**	**Factor 3 (n=2)**
Intermediate / Medium (MUET Band 3–4)	5	1	2
Advanced / High (MUET Band 5–6 / IELTS)	9	4	0

Although participants in this study had diverse first-language (L1) backgrounds, L1 background did not emerge as a primary basis for factor membership. In other words, learners with different L1s were not clearly separated into different Q-factors. However, this does not mean that L1 was irrelevant. Rather, interview data suggest that L1 functioned as a mediating linguistic resource across factors, helping learners ask questions, compare expressions across languages, interpret pragmatic nuances, and manage accent-related concerns when interacting with AI. For instance, P12 (L1: Chinese) explained that when expressing complex ideas or emotions, she typically organizes her thoughts in her L1 first and then asks AI to help translate them into English, in order to avoid vague or inaccurate expressions. P16 similarly noted that he frequently uses L1 to ask questions and then converts AI's responses into English output, which he believes is comprehensible. In addition, some participants employed L1 to cope with accent and pragmatic concerns. P19 mentioned that when uncertain about an expression, he first describes his intended meaning in L1 and then asks AI to provide an appropriate English formulation to check naturalness. Likewise, P03 described English as "one of my mother tongues" and reported naturally mixing English and Indonesian when interacting with AI, particularly when interpreting films, slang, or culturally embedded expressions through cross-linguistic comparison. Taken together, learners do not treat AI as a monolingual tool, but rather embed it within their existing multilingual resource systems, flexibly mobilizing L1 to support comprehension, evaluation, and spoken English production.

5 Discussion

5.1 AI use logics reflected across factors

From a theoretical perspective, the three factors identified in this study can be understood as distinct shared viewpoints rather than as simple differences in the degree of AI use in speaking, which is consistent with the logic of Q-methodology (Brown, 1993; Watts & Stenner, 2005). Interpreted through UTAUT2, these viewpoints show that learners combine acceptance-related beliefs such as usefulness, ease of use, and facilitating conditions with concerns that are less fully captured by the model, including emotional authenticity, privacy, and overreliance. This supports earlier Q-methodology work showing that learners' perception of AI-mediated language learning is structured, mixed, and not reducible to isolated variables (e.g., Guo & Xia, 2025). Unlike studies that dichotomize learners into “active” versus “inactive” users (e.g., An et al., 2023; Wang, 2024), this study shows that even learners who use AI frequently or competently may develop fundamentally different orientations toward AI use.

Specifically, learners aligned with the performance-oriented factor primarily interpret AI’s value in terms of task completion and performance enhancement, treating it as an instrumental resource for improving speaking quality and optimizing language output. This orientation is consistent with IDLL literature on the role of AI in enhancing efficiency and supporting self-directed learning (Liu et al., 2024). However, this study revealed that these learners clearly distinguish between preparatory support and authentic expression, while maintaining vigilance against overreliance.

In contrast, the emotion-driven factor does not represent a rejection of AI’s functional value, but rather highlights learners’ heightened sensitivity to interactional experience, affective feedback, and emotional comfort (Chassignol et al., 2018). While previous research emphasized AI’s non-judgmental nature as an advantage in reducing speaking anxiety (Yang & Li, 2024), the present study showed that this advantage is not uniformly perceived. Some learners remain highly alert to AI’s emotional “flatness,” pragmatic unnaturalness, and lack of authentic responsiveness, which constrains AI’s role in speaking practice (Kohnke & Moorhouse, 2025). This factor illustrates that emotion does not simply facilitate technology use, but it can also function as a critical lens through which learners reassess AI’s appropriateness.

The third factor is better understood as reflecting a convenience- and accessibility-oriented viewpoint. From a UTAUT2 perspective, Factor 3 appears to reflect strong performance expectancy, effort expectancy, and facilitating conditions, especially in relation to immediate availability and ease of use. However, this viewpoint includes selective caution about specific limitations, particularly pronunciation support, rather than broad concern about linguistic unnaturalness or cognitive overdependence. This tension has not always been explicitly foregrounded in prior AI acceptance research, which has often emphasized general predictors of use in areas such as academic writing (e.g., Wang, 2024). In the present study, however, it becomes more visible in AI-mediated speaking practice. This is likely because oral production involves immediacy, exposure, and continuous monitoring. These findings suggest that AI’s compensatory function in speaking practice is not purely empowering, but may also amplify feelings of insecurity and dependence. Overall, the three factors do not form a linear continuum from low to high AI use, but instead represent qualitatively different logics of practice.

5.2 English proficiency and L1 background as contextual conditions

From the results, self-reported English proficiency showed a clearer relationship with factor alignment than L1 background. L1 did not systematically differentiate the extracted factors, but it still played an important mediating role in learners’ AI use across those factors. More specifically, learners drew on their L1 as a multilingual support resource for prompting AI, checking meanings, comparing expressions, and clarifying culturally or pragmatically appropriate English use. Thus, in this study, L1 is better understood as a cross-factor resource shaping practice rather than as a variable determining viewpoint membership (Liu et al., 2024). This interpretation should be treated cautiously, however, because proficiency was based on self-report and was not externally validated in the present study.

A key difference is that advanced learners in this study did not simply use AI “more” or show greater willingness to adopt it; rather, they demonstrated a stronger capacity to delimit AI use to specific purposes and stages, while actively guarding against overdependence. This aligns with recent discussions on AI literacy and metacognitive AI use strategies. Prior studies have often reported that intermediate-level learners’ performance expectancy positively predicts usage intention (e.g., An et al., 2023; Wang, 2024). However, the present study suggests that for these learners, high performance expectancy is frequently bundled with intensified anxiety about language quality and cognitive erosion, resulting in oscillation between reliance on AI for efficiency compensation and fear of homogenization or weakened thinking. This pattern has not always been made explicit in prior studies on AI acceptance because studies have often focused more on general predictors of use in domains such as academic writing than on the coexistence of efficiency-related benefits and anxiety-related concerns (e.g., Wang, 2024). The present study makes this tension more visible in AI-mediated speaking practice.

Regarding multilingual background, the present findings differ from Guo and Xia's (2025) study on AI-IDML, which reported language-specific motivational patterns across English and Japanese AI-IDLL. The current study revealed that within the same English-speaking practice, learners' L1 backgrounds did not directly divide factors, but instead manifested in the strategic use of L1 resources to mediate AI output. This finding supports Yu et al.'s (2025) argument that AI use in multilingual contexts is highly translingual and strategic, and further suggests that multilingual identity operates more as a regulator of usage practices than as a determinant of stance.

The findings of the present study have more specific pedagogical and design implications. For Factor 1 learners, AI tools should support goal-oriented practice through features such as task rehearsal, pronunciation feedback, vocabulary expansion, and progress tracking, while teachers should encourage critical evaluation to prevent overreliance. For Factor 2 learners, AI design should emphasize emotional credibility, natural voice, culturally appropriate interaction, and transparent privacy settings, ideally combined with human-supported practice. For Factor 3 learners, AI tools should provide structured scaffolding, such as guided prompts, simplified feedback, and "think-first" support, so that efficiency benefits do not increase anxiety or dependence. These findings suggest that learner-centred AI-speaking support should vary according to learners' dominant orientations rather than assume uniform needs.

6 Conclusion

This study used Q-methodology to examine multilingual Malaysian undergraduates' perception of AI-IDESP. Three factors emerged: performance-oriented strategic use, emotion- and risk-sensitive caution, and efficiency-driven reliance with anxiety. The study revealed that AI-mediated speaking practice varies according to learners' priorities, emotions, and linguistic confidence. Specifically, English proficiency was related to alignment with different viewpoints. Higher-proficiency learners viewed AI as a performance-enhancing tool while remaining critically aware of its limitations. Intermediate learners relied on AI for accessibility and emotional safety but expressed concerns about dependence and authenticity. Thus, AI-IDESP is not a uniform solution; its perceived value and risks vary by proficiency and learner needs.

The study advances AI-IDLE research by capturing holistic viewpoints rather than isolated acceptance variables. Theoretically, the findings refine the role of UTAUT2 in AI-mediated language learning. While the factors partly reflect established constructs such as performance expectancy, facilitating conditions, and habit, they also reveal dimensions that are less visible in UTAUT2, including emotional authenticity, privacy concerns, speaking anxiety, and fears of cognitive overreliance. In this sense, the study suggests that understanding AI-IDESP in multilingual contexts requires combining technology-acceptance perspectives with affective and cognitive theories of language use. Methodologically, the study highlights the value of integrating Q-sorts with interviews to reveal structured patterns and lived experiences.

Pedagogically, AI speaking tools should be implemented with attention to learner diversity: advanced learners may use AI as a supplementary performance aid, while intermediate learners may need guidance to prevent overreliance and support engagement with authentic interaction. Pedagogically, the findings suggest that AI speaking tools should be implemented with greater sensitivity to learner differences. For performance-oriented learners, AI may function best as a task-rehearsal and feedback tool; for emotionally sensitive learners, it should be supplemented with more authentic and relationally supportive interaction. For learners who depend on AI for efficiency but experience anxiety, more structured scaffolding is needed to reduce cognitive overload and prevent overreliance. In design terms, this means that learner-centred AI tools should incorporate not only feedback and accessibility, but also privacy controls, culturally appropriate interaction, natural voice options, and graduated support for independent speaking. Future research could employ longitudinal designs or examine other multilingual contexts to track evolving orientations toward AI. Finally, students' proficiency-related interpretation in this study should be viewed with caution, as English proficiency in the study was self-reported rather than independently validated. Future research could incorporate standardized proficiency measures to examine these patterns more robustly.

Ethical Statement

Under the institution's guidelines for low-risk educational research, this type of study does not require formal ethics review. Accordingly, no formal ethics approval or exemption number was issued. All participants provided informed consent prior to participation. Participation was voluntary, anonymity was maintained, and participants were informed of their right to withdraw at any time without penalty.

GenAI Use Disclosure Statement

The authors declare no use of Generative AI.

References

Agyei, C., & Razi, Ö. (2022). The effect of extended UTAUT model on EFLs' adaptation to flipped classroom. *Education and Information Technologies, 27*(2), 1865-1882.

An, X., Chai, C. S., Li, Y., Zhou, Y., & Yang, B. (2023). Modeling students' perceptions of artificial intelligence assisted language learning. *Computer Assisted Language Learning*, *38*(5-6), 987-1008. https://doi.org/10.1080/09588221.2023.2246519

Benson, P. (2011). Language learning and teaching beyond the classroom: An introduction to the field. In *Beyond the language classroom* (pp. 7-16). Springer.

Brown, S. R. (1993). A Primer on Q Methodology. *Operant subjectivity*, *16(3/4)*(3/4).

Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial Intelligence trends in education: a narrative overview. *Procedia computer science*, *136:* 16-24.

Dewaele, J.-M., & MacIntyre, P. D. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Studies in Second Language Learning and Teaching*, *4*(2): 237-274. https://doi.org/10.14746/ssllt.2014.4.2.5

Dewaele, J.-M., Witney, J., Saito, K., & Dewaele, L. (2018). Foreign language enjoyment and anxiety: The effect of teacher and learner variables. *Language teaching research*, *22*(6): 676-697.

Dressman, M., & Sadler, R. W. (2020). *The handbook of informal language learning*. John Wiley & Sons.

Fang, F., Meng, Y., Tang, L., & Cui, Y. (2025). The Impact of Informal Digital Learning of English (IDLE) on EFL Learners' Engagement: Mediating Roles of Flow, Online Self-Efficacy, and Behavioral Intention. *Behav Sci (Basel)*, *15*(7). https://doi.org/10.3390/bs15070851

Gill, S. K. (2013). *Language policy challenges in multi-ethnic Malaysia* (Vol. 8). Springer Science & Business Media.

Guo, S., & Xia, M. (2025). Exploring Learners' Perceptions of AI-mediated informal digital Multilingual Learning: A Q-Methodology Approach. *International Journal of Applied Linguistics*. https://doi.org/10.1111/ijal.12815

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign Language Classroom Anxiety. *The Modern language journal,* , *70*(2):125-132.

Irie, K., Ryan, S., & Mercer, S. (2018). Using Q methodology to investigate pre-service EFL teachers' mindsets about teaching competences. *Studies in Second Language Learning and Teaching*, *8*(3):575-598. https://doi.org/10.14746/ssllt.2018.8.3.3

Kohnke, L., & Moorhouse, B. L. (2025). Enhancing the emotional aspects of language education through generative artificial intelligence (GenAI): A qualitative investigation. *Computers in Human Behavior*, *167*, 108600.

Lee, J. S. (2018). Informal digital learning of english:the case of korean university students. *Doctoral dissertation, University of Illinois at Urbana-Champaign*. https://hdl.handle.net/2142/101354

Lee, J. S. (2019). EFL students' views of willingness to communicate in the extramural digital context. *Computer Assisted Language Learning*, *32*(7): 692-712. https://doi.org/https://doi.org/10.1080/09588221.2018.1535509

Lee, J. S., & Lee, K. (2021). The role of informal digital learning of English and L2 motivational self system in foreign language enjoyment. *British Journal of Educational Technology*, *52*(1): 358-373. https://doi.org/10.1111/bjet.12955

Liu, G. L., Darvin, R., & Ma, C. (2025). Exploring AI-mediated informal digital learning of English (AI-IDLE): a mixed-method investigation of Chinese EFL learners' AI adoption and experiences. *Computer Assisted Language Learning*, *38*(7): 1632-1660. https://doi.org/10.1080/09588221.2024.2310288

Liu, G. L., Zhao, X., & Yang, B. (2024). The predictive effects of motivation, enjoyment, and self-efficacy on informal digital learning of LOTE: Evidence from French and German learners in China. *System*, *126*: 1-15. https://doi.org/10.1016/j.system.2024.103504

Liu, H. (2019). *New insights into language anxiety: Theory, research and educational implications* (Vol. 73). Oxford University Press UK.

Liu, X. J., Wang, J., & Zou, B. (2025). Evaluating an AI speaking assessment tool: Score accuracy, perceived validity, and oral peer feedback as feedback enhancement. J*ournal of English for Academic Purposes*, *75*, 1-14.

MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The modern language journal*, *82*(4): 545-562.

Nguyen, C. H. (2026). "I know I should, but I don't": A case study of Vietnamese students' reluctance to engage in IDLE. *International Journal of TESOL Studies, 8*(2), 73–90. https://doi.org/10.58304/ijts.250911

Richards, J. C. (2015). The Changing Face of Language Learning: Learning Beyond the Classroom. *RELC Journal*, *46*(1): 5-22. https://doi.org/10.1177/0033688214561621

Soyoof, A., Reynolds, B. L., Vazquez-Calvo, B., & McLay, K. (2021). Informal digital learning of English (IDLE): a scoping review of what has been done and a look towards what is to come. *Computer Assisted Language Learning*, *36*(4): 608-640. https://doi.org/10.1080/09588221.2021.1936562

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, *27*(3): 425-478. https://doi.org/https://doi.org/10.2307/30036540

Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS quarterly*, 36(1): 157-178.

Wang, Q. (2024). EFL learners' motivation and acceptance of using large language models in English academic writing: an extension of the UTAUT model. *Front Psychol*, *15*, 1514545. https://doi.org/10.3389/fpsyg.2024.1514545

Watts, S., & Stenner, P. (2005). The subjective experience of partnership love: AQ methodological study. *British journal of social psychology, 44*(1), 85-107.

Watts, S., & Stenner, P. (2008). Doing Q methodology: theory, method and interpretation. *Qualitative Research in Psychology*, *2:* 67-91. https://doi.org/10.1191/1478088705qp022oa

Xing, C., & Saeed, M. A. (2025). A Systematic review on artificial intelligence (AI) technologies in ESL/EFL speaking skills. I*nternational Journal of TESOL Studies*, *8*(2), 240-270. https://doi.org/10.58304/ijts.250908

Yang, L., & Li, R. (2024). ChatGPT for L2 learning: Current status and implications. *System*, *124*: 1-14. https://doi.org/10.1016/j.system.2024.103351

Yu, H., Guo, Y., Yang, H., Zhang, W., & Dong, Y. (2025). Can ChatGPT revolutionize language learning? Unveiling the power of AI in multilingual education through user insights and pedagogical impact. *European Journal of Education*, *60*(1), e12749. https://doi.org/https://doi.org/10.1111/ejed.12749

Zhang, Y., & Liu, G. (2022). Revisiting informal digital learning of English (IDLE): a structural equation modeling approach in a university EFL context. *Computer Assisted Language Learning*, *37*(7): 1904-1936. https://doi.org/10.1080/09588221.2022.2134424

Zhang, Y., Lucas, M., Bem-haja, P., & Pedro, L. (2025). AI versus human-generated voices and avatars: rethinking user engagement and cognitive load. *Education and Information Technologies*, 1-20.

Zou, M., Soyoof, A., Teng, M. F., & Chen, H. (2025). A skill-specific perspective on AI-mediated informal digital learning of English (AI-IDLE): Examining the contributing roles of L2 writing motivation and enjoyment. Language Teaching Research. Advance online publication. https://doi.org/10.1177/13621688251375547

Zou, M., Teng, M. F., Soyoof, A., & He, X. (2026). Informal digital learning of English (IDLE) as form-focused and meaning-focused activities: Refining its measurement and examining its predictive role in L2 achievement and confidence. *International Journal of Applied Linguistics, 36*, 393–408. https://doi.org/10.1111/ijal.12781

Li Rong is a Master student in English Language Studies at the Department of English Language, Faculty of Language & Linguistics, Universiti Malaya.

Murad Abdu Saeed is a Senior Lecturer at the Faculty of Languages & Linguistics, Universiti Malaya, Kuala Lumpur. His areas of research interest include Educational technology, feedback, engagement, feedback literacy, & collaborative writing.