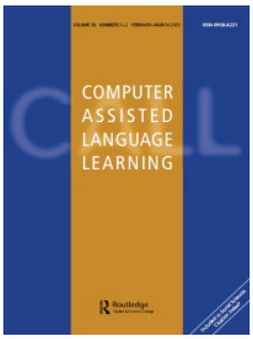

Computer Assisted Language Learning

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/ncal20

The role of design thinking skills in artificial-intelligence language learning (DEAILL) in shaping language learners' L2 grit: the mediator and moderator role of artificial intelligence L2 motivational self-system

Amir Reza Rahimi & Ana Sevilla-Pavón

To cite this article: Amir Reza Rahimi & Ana Sevilla-Pavón (09 Apr 2025): The role of design thinking skills in artificial-intelligence language learning (DEAILL) in shaping language learners' L2 grit: the mediator and moderator role of artificial intelligence L2 motivational self-system, Computer Assisted Language Learning, DOI: 10.1080/09588221.2025.2477710

To link to this article: https://doi.org/10.1080/09588221.2025.2477710

COMPUTER ASSISTED LANGUAGE LEARNING
https://doi.org/10.1080/09588221.2025.2477710

Routledge
Taylor & Francis Group

The role of design thinking skills in artificial-intelligence language learning (DEAILL) in shaping language learners' L2 grit: the mediator and moderator role of artificial intelligence L2 motivational self-system

Amir Reza Rahimi[a] and Ana Sevilla-Pavón[b]

[a]Language, Literatures and Cultures and its Applications, University of Valencia, Valencia, Spain; [b]IULMA/Faculty of Philology, Translation and Communication, IULMA/Universitat de València, Valencia, Spain

ABSTRACT

Among the most effective ways to develop individuals' skills and competences in the twenty first century is to use Design Thinking (DT) as a problem-solving approach. Several fields of study have employed this approach to facilitate students' problem-solving skills through Information and Communications Technologies (ICT). Nevertheless, this skill has been glaringly overlooked in Computer-assisted Language Learning (CALL). For this sake, the researchers developed the Design Thinking Skills in Artificial Intelligence Language Learning (DEAILL) scale to address this gap in Artificial Intelligence Language Learning (AILL) and examined its role in shaping Artificial Intelligence L2 motivational self-system (AIL2MSS) and L2 grit among 92 Spanish English as a Foreign Language (EFL) students. Having validated the factorial structures of DEAILL, AIL2MSS and L2 grit in the study context, the Partial Least Structural Equation Modeling (PLS-SEM) showed that the more often language learners collected AI feedback based on their areas of needs and applied it to their target learning context, the more positive their future image and current L2 identity in AILL became. Additionally, the sign of authenticity gap was observed in this context, demonstrating participants' perception of AILL and DEAILL as more authentic contexts and skills compared with their previous learning experiences and skills. This mediated the relationship between learners' DEAILL skills and their L2 grit to devote more interest and persistence to AILL problem-solving. Consequently, the study proposes a new theoretical framework for CALL and the Common European Framework of Reference for Languages (CEFR) and recommends that language teachers prioritize the development of learners' problem-solving skills rather than learning outcomes. Furthermore, it is advisable for pedagogical

ARTICLE HISTORY
Received 2 March 2024
Accepted 5 March 2025

KEYWORDS
21st-century digital skills; 21st-century digital competence; artificial intelligence in education; problem-solving in education; complex dynamic systems

CONTACT Ana Sevilla-Pavón ana.m.sevilla@uv.es IULMA/Universitat de València, IULMA/Faculty of Philology, Translation and Communication, 32 Blasco Ibáñez Avenue, 46010 Valencia, Spain.

experts to offer targeted teacher training programmes that equip language instructors with the skills and knowledge to foster DEAILL skills through CALL tools. Such professional development would ensure that teachers are prepared to integrate technology effectively into their language teaching. Additionally, including DEAILL competence as a mandatory criterion for language instructors could further enhance the quality of technology-enhanced language education.

Introduction

With the development of technology and artificial intelligence (AI) at its core, human life has been profoundly impacted in many areas, particularly in education. This has led several organizations, including Assessment and Teaching of Twenty-First Century Skills (ATC21S), the Organization for Economic Co-operation and Development (OECD), the European Union (EU), the Computer Science Teachers Association (CSTA), the International Society for Technology in Education (ISTE), and the International Association for the Evaluation of Educational Achievement (IEA) to change the process of teaching with ICT and AI to problem-solving and lifelong learning (Rahimi, 2024b). This requires both teachers and learners to become innovative problem solvers by cultivating 21st-century digital skills through ICT and AI (Rahimi, 2024b; Rahimi & Sevilla-Pavón, 2025b). In this regard, the Digital Education a Action Plan 2021–2027 has emphasized the importance of students possessing a sense of competence and skills in order to integrate 21st-century digital skills (Outeda, 2024), as it 'involves the confident, critical and responsible use of, and engagement with, digital technologies for learning, at work, and for participation in society' (The Council of Europe, 2018, as cited in Rahimi, 2024a, p. 3). In this line, Rahimi (2024b), and Rahimi and Mosalli (2024) encouraged English Language Teaching (ELT) and Computer-assisted Language Learning (CALL) researchers to take a broader view of CALL materials and AI usage in language teaching and learning, going beyond language teaching skills and sub-skills and cultivating 21st-century digital skills.

The above recommendations have led to the development of curricula in most subjects, particularly in science, technology, engineering, and mathematics (STEM) based on constructivist learning, problem-solving, as well as 21st-century digital skills (Rahimi, 2024b). A critical 21st-century digital skill is Design Thinking(DT). It is characterized by analytical, algorithmic and creative processes where individuals have the opportunity to experiment, create, gather feedback and redesign their target outcome with high validity (Liu & Li, 2023). Due to this, over time, DT has

become an integral part of STEM education to assist students in achieving 21st-century digital competences and skills to prepare future generations for the post-modern era by living, learning and solving problems with ICT and AI. However, it is important to note that, as highlighted by Brown and Katz (2011), DT is not a skill confined to STEM education. Rather, DT is a skill that connects different fields, as it offers a 'structured method that encourages participants to empathize with users and their contexts, thereby fostering the generation of innovative solutions' (as cited in Leem & Lee, 2024, p. 2). Accordingly, with the development of ICT and AI, most fields, such as psychology, law (Hews et al., 2023), and business (Chouyluam et al., 2021), have integrated DT into their curriculum. There is, however, a tendency for CALL to focus primarily on teaching language skills or subskills. As a result, it often falls short of integrating 21st-century digital skills, particularly Digital Thinking (DT) skills. These skills are crucial for learners to adapt their language learning processes to the current technological landscape, where tools such as large language models (LLMs), intelligent chatbots, automatic writing correction, machine translation and adaptive recommendation systems have recently been incorporated into Artificial Intelligence Language Learning (AILL).

Integrating these tools can help learners not only enhance their language skills but also prepare them to think more creatively and innovatively when solving complex and open-ended problems within the contexts of learning, working, and living with artificial intelligence. Since the advent of AI in human life, education has increasingly provided individualized language learning and problem-solving opportunities (Rahimi & Sevilla-Pavón, 2025b). However, this integration has raised concerns about inaccuracies and hallucinations in AI-generated content (Crompton & Burke, 2024). To address these challenges, an empathetic and practical DT approach, focused on solving real-world problems through collaboration, could be key. This approach, however, has yet to be explored in depth within applied linguistics and CALL, despite its potential to mitigate the aforementioned issues. Nevertheless, it is important to bear in mind the need for students to be critical when dealing with the information provided by chatbots as part of their personalized language learning experience. This highlights the importance of fostering DT in language education, particularly in CALL, as stressed by Rahimi (2024b) to cover 21st-century digital skills in the field of CALL nowadays. More importantly, CALL experts suggest that CALL will continue to advance as technology advances (Colpaert, 2020) and, accordingly, the aforementioned gaps are now addressed as we go beyond language teaching skills by implement problem-solving solutions and in doing so, validate DT skills within CALL and AILL.

The process of learning a foreign language requires exposure to a wider context, skills and external factors, including the learning environment, the availability of existing resources and exposure to training languages in a natural setting. It is thus essential to determine whether the instruction will take place remotely, in-person or in a hybrid format. This last factor has become important recently because of the huge growth and development of chatbots and the use of AI in language learning, which have led to personalized learning experiences which are mostly provided by chatbots and ChatGPT (Rahimi & Sevilla-Pavón, 2024). Thus, their use has become standard in AILL. Due to the introduction of AI into the classroom, both teachers and students face new challenges and requirements, such as reaching accurate and reliable responses from AI assistants. These in turn have contributed to the integration of 21st-century digital skills into the learning process, and teachers are now encouraged to help students develop those skills in order to make appropriate usage of AI and ICT, especially when learning foreign languages (Rahimi, 2024b). Nevertheless, the successful cultivation of any skills in L2 learning appears to be dependent on the development of particular psychological dispositions. The motivation of learners and their grit are among those factors, as recently reported by Paradowski & Jelińska, 2023, and parallel to the advances of ICT and AI, language learners' psychological aspects have also become more sophisticated, which has led to the development of new terminologies such as online motivational self-systems (Rahimi & Cheraghi, 2022), online self-regulation (Zheng et al., 2018), and digital self-authenticity (Rahimi, 2023; Smith et al., 2020) in this field. Another concept that recently attracted psycholinguistic researchers is L2 grit, defined as 'perseverance and passion for long-term goals' (Duckworth et al., 2007, p. 1087). In recent years, the study of L2 grit and language learners' motivation has gained popularity among L2 researchers in both online (Rahimi & Sevilla-Pavón, 2025a) and blended language learning contexts. However, further research is necessary to understand the integration of these new psychological factors into other areas of CALL, particularly AILL. This is because dynamic complex systems researchers have found that psychological factors like motivation, self-regulation and L2 grit are interconnected and depend on each other in every language learning context (Mercer, 2018). Furthermore, these connections have been looked into with regard to language learners' skills and subskills (Paradowski & Jelińska, 2023). However, the lack of examination of their intercorrelation with 21st-century digital skills, specifically Design Thinking skills in Arficial Intelligence Language Learning, which might provide a new conceptual framework for SLA and CALL, highlights a gap that the present study could potentially fill.

2. Literature review

2.1. Design thinking skills in artificial-intelligence language learning (DEAILL)

Peter Rowe introduced the concept of design thinking in his 1991 book *Design Thinking* (Rowe, 1991), defining DT a rational set of human actions for improvement (Kimbell, 2011). DT was initially applied in engineering, management and medicine (Simon, 2019). However, as time passed and ICT became more integrated, it became a necessary skill for a wider variety of fields. In this regard, Brown (2008) describes DT as an approach to technology development that takes into account individuals' needs, behaviors and the feasibility of the technology. It is also known as the method that designers use to come up with solutions to poorly defined problems (Micheli et al., 2018) and it has been viewed as a tridimensional process including inspiration, ideation and implementation (Brown & Wyatt, 2010): in the first stage, the individual searches for a solution, then develops their ideas and executes them. The Stanford D. School (2010) subsequently developed a conceptual framework of DT for general education, comprising five steps: empathize, define, ideate, prototype and test. Educational contexts often use this framework to enhance learners' DT skills. With the increasing complexity of technology, the human-centered perspective of DT offered perspectives connected to its use to find solutions to some of the more complex technological problems associated with current human needs. According to Dym et al. (2005), DT employs divergent-convergent questioning, with convergent attributes for deep reasoning questions and divergent attributes for generative design questions. The capacity of DT to facilitate participant tolerance for ambiguity and failure is the basis for its widespread use in education courses that address real-world problems (Leem & Lee, 2024). Since CALL and AILL aim to integrate technology to provide language learners with real-world language learning experiences, the integration of new-generation of LLM has brought more opportunities in this regard. However, it is important to bear in mind the risks inherent to such personalized learning environments, since AI might provide inaccurate and hallucinatory information (Crompton & Burke, 2024). This underscores the need to cultivate design thinking skills among language learners, as its human process structure in artificial intelligence education has been shown to improve students' learning process and critical thinking. Furthermore, it has also been reported as helpful for obtaining the best possible answers and solutions to address problems and challenges, especially in the context of AI education (Lin & Chang, 2024; Sreenivasan & Suresh, 2024; Staub et al., 2023). However, this process has not been explored in the context of AILL. Since it might

be time-consuming, it requires high levels of L2 motivation, persistence of effort and constancy of interest. Therefore, we have taken a significant step forward in both CALL and AILL by operationalizing DEAILL as follows:

- Define: language learners' identification of the areas of language needs and problems they want to remedy with the assistance of AI.
- Ideate: language learners' brainstorming and the collection of feedback and assistance provided by AI.
- Prototype: language learners' transformation and application of AI feedback and guidance to their language-learning context.
- Test: language learners' awareness of AI guidance reliability through getting feedback from their language teachers or peers.

2.2. General grit and L2 grit

Despite the common belief that talent is the key to success, Duckworth et al. (2007) suggest that successful individuals share grit as well as talent. Grit means staying positive despite setbacks, maintaining interest, and overcoming challenges. Duckworth et al. (2007) describe two components of grit: consistency of interests, which is a long-lasting passion regardless of failure, disappointment, or challenges over a sustained period; and persistency of effort, which means perseverance with long-term pursuits that require continuous energy investment. These authors designed and validated the original grit scale (Grit-O; Duckworth et al., 2007) and its short version (Grit-S; Duckworth & Quinn, 2009), and the aforementioned universal grit components have been validated, translated into other languages (e.g. Chinese, Turkish, Spanish), and utilized in several research works. Many studies have demonstrated that grit is associated with successful completion of demanding training programs, higher academic performance, improved teacher effectiveness, creative thinking and the ability to cope with changes (Roberts, 2009). As a result, many SLA studies have utilized Grit-O and Grit-S. For clarification, Lake and D. Da Silva (2013) conducted a pioneering study investigating the relationship between grit and a range of positive personality dispositions among female Japanese ESL students using the original grit scale. Based on L2 motivational measures and positive self-measures, they found that grittier students invested more effort in EFL learning and had greater curiosity, hope, subjective happiness and flourishing. Subsequently, Changlek and Palanukulwong (2015) employed the original grit scale and discovered a positive correlation between motivation and grit among Indian EFL learners, whereas anxiety showed a negative correlation with grit. In another study, looked into Chinese

EFL students' grit using the original grit scale and discovered that PE predicted both the intensity of motivation and the persistence of effort in learning an additional language. Despite this, recent research works have questioned the predictive power of general grit. To clarify, in their meta-analysis, Credé et al. (2017) reported that grit does not contribute statistically significantly to cognitive and cognitive constructs but is not negligible in high-stakes situations. In other words, after controlling for the contribution of conscientiousness, grit failed to explain any variance in academic performance, and it explained very little variance in high school performance. The same method for calculating grit was used by Lam and Zhou (2022), who declared that the correlation between grit levels and academic success was weak to moderate and that PE made a greater contribution than CI, suggesting that the two constructs are distinct from one another. According to these reports, the non-cognitive construct of grit is less useful than conscientiousness, representing what Credé et al (2017) termed 'old wine in new bottles', which is particularly pertinent when perseverance does not contribute a unique contribution or explain a modest unique variance. As a result of these reports, debates about the usefulness of grit as a construct continued, which suggests that it is necessary to revisit this non-cognitive trait in order to resolve these concerns (Credé et al., 2017). Having addressed these issues in second language acquisition, Teimouri et al. (2020) proposed that grit, as with other psychological factors such as language learning motivation and self-regulation, has its own domains and thus should not be tapped into broadly using general-domain measures (Teimouri et al., 2020). Therefore, a language-specific grit scale was developed and it was reported that it was highly correlated with language learners' motivation and achievement in comparison with general grit, providing supportive evidence for the construct and showing the validity of L2 grit measures. Studies have also shown that there was either a weak association between domain-general grit and L2 achievement (e.g. Wei et al., 2019) or no association at all (e.g. Yamashita, 2018). In contrast, the L2 grit domain had a high predictive power in language learners' achievement (Alamer, 2021). In a more recent study, Li and Yang (2023) combined domain-general grit with L2 grit, reporting that general grit lost all predictive power while L2 domain remained strong. As a result, L2 grit appears to have a high degree of predictive power, which has recently been applied and explored in the fields of CALL. Zarrinabadi et al. (2022), in their study, highlighted the crucial role L2 grit played in connection to persistence of effort and the CALL surplus value of CALL. Similarly, in their recent study on Virtual Exchange, Rahimi and Sevilla-Pavón (2024) highlighted the key role of learners' L2 grit in moderating and mediating the relationships between learners' online

motivation and online self-regulation in VE. However, its role in other CALL subfields, such as AILL in combination with other psychological or new skills, has not been explored yet.

2.3. Artificial intelligence L2 motivational self-system

A thriving area of L2 studies has always been motivation, particularly Dörnyei's (2005) L2 motivational self-system. The notion of English as a Lingua Franca (ELF) in the EFL context has gained traction due to globalization over the past few years and thanks to Dörnyei's (2009) groundbreaking finding that actual language use is what shapes students' motivation rather than integrativeness in EFL contexts, since students do not usually have direct contact with the target language community in such contexts. Hence, the L2MSS theory replaces the traditional socio-psychological model of language learning motivation by emphasizing three main aspects: the ideal L2 self, the ought-to L2 self, and the L2 learning experience. ICT have also helped to advance this theory, allowing L2 scholars to adapt or extend its constructs in a variety of ways. For clarification, Zheng et al. (2018) developed online language learning motivation (OLLM), which encompasses several constructs, including instrumentality-promotion (IPO), cultural interest, instrumentality prevention (IPR), other' expectations, and online language learning experiences. Several studies have already demonstrated the importance of IPO and IPR in online language learning. Zheng et al. (2018) highlighted the positive impact that Chinese EFL learners' IPR had on shaping their online self-regulation. Additionally, Rahimi and Cheraghi (2022) found that Iranian English learners' IPO and IPR significantly influenced their task strategies, help-seeking and environment structuring within an LMOOC, resulting in their LMOOC's course retention. Rahimi (2023) also highlighted the predictive power of both IPO and IPR in predicting language learners' behavioral intentions to learn in LMOOCs in the future. Thus, the positive effects of both IPR and IPO in shaping language learners' behaviors have already been highlighted in online language and LMOOC. In our current study, we investigate them in AILL to gain a more profound comprehension of their function in a different CALL sub-field.

Henry (2013) also reported that language learners exhibited a high level of motivation in digital language learning, while their motivation and language learning effort in their face-to-face language learning classroom declined. Thus, he found an *authenticity gap* between these two contexts, as the digital learning environment was more tailored to their needs. Furthermore, Henry and Cliffordson (2017) observed this authenticity gap in Swedish EFL learners and reported similar findings. To

extend their work, Smith et al. (2020) replicated their study in Chinese online language learning and reported that digital self-authenticity positively predicted their participants' online language learning effort. Continuing on their path, Rahimi (2023) and Rahimi and Mosalli (2024) confirmed the validity of digital self-authenticity as an L2MSS factor in the Iranian EFL context and extended their previous findings to shape the attitudes of language learners in LMOOCs and emergency remote language learning. Thus, we will now investigate the role of digital self-authenticity in another CALL context and explore its role in shaping language learners' L2 grit.

Furthermore, recent systematic and meta-analyses have found L2MSS to be one of the most predictive theoretical frameworks. Al-Hoorie (2018) published a meta-analysis of L2MSS, finding that more than 32 studies applied this theoretical framework to 30,000 examples from a variety of linguistic situations (e.g. You & Dörnyei, 2015), reporting its predictive power for language learning motivation. In this vein, Al-Hoorie (2018) published a meta-analysis of L2MSS, reporting its predictive power for language learning motivation in 32 studies applying this theoretical framework to 30,000 examples from a variety of linguistic situations (e.g. You & Dörnyei, 2015). The synthetic report by Yousefi and Mahmoodi (2022) affirmed that the L2MSS is important for understanding language learners' motivation to learn a language, as 17 published studies have already applied it to 18,832 participants. Because of these positive reports, the researchers chose and adopted the following L2MSS constructs based on previous conceptual frameworks (Rahimi, 2023; Smith et al., 2020), and defined them as follows:

- Instrumentality-promotion (IPO): language learners' personal future images or personal visions of their future goals, such as achieving native-like pronunciation or solving problems with AI through language learning.
- Instrumentality-prevention (IPR): language learners (with saxon genitive) recognition of the need to learn and solve their language learning challenges with AI, such as passing a language and culture course in university to advance to the next semester.
- Current L2 self (CU): students' self-descriptions about their current language competence and abilities to use, learn, and solve their language learning problems with AI.
- Authenticity gap (AUT): comparison of language learners' motivation, attitudes and personal objectives in learning and addressing language challenges, using AI versus traditional language learning methods.

2.4. Literature review and gaps

Researchers have widely adopted the DT due to its multiple benefits in improving learners' 21st-century problem-solving skills by promoting the acquisition of innovative and complex problem-solving skills, mainly through AI (Calavia et al., 2023; Yang et al., 2022). According to these studies, learners who acquire DT can transform their imaginations, creative concepts and social experiences into practical applications through experimentation to enhance their potential for practically resolving problems in natural settings (Calavia et al., 2023; Yang et al., 2022). For clarification, Tsai et al. (2023) applied self-determination (SDT) as learners' core and basic psychological needs and designed a constructivist and problem-solving course for learners to solve problems through DT. The t-test results revealed that DT increased learners' motivation significantly. Similarly, Tsai and Wang (2020) investigated its role in enhancing STEM students' self-efficacy. They determined that recognizing problems, brainstorming solutions, and applying them led to increased self-efficacy in STEM learners. Similarly, in her qualitative study, Balakrishnan (2021) found that DT motivated learners to be creative and propose and develop practical, innovative problem-solving strategies in a natural setting.

DT skills have also been examined by language researchers in their SLA studies, particularly in the context of traditional language learning. For instance, Opincāne and Laganovska (2023) reported on the effectiveness of DT, as language learners were able to collaborate and participate in discussions of simulated professional and real-life issues within their language learning context and in real-world situations. Another example can be found in Frolova and Aleshchanova's (2021) experimental study involving the cultivation of DT skills by language teachers in an experiential group. They found that the performance on intellectual modeling, research activities and subject communication in the target language was superior among the students from the experimental group. In the same year, Cleminson and Cowie (2021) also reported that DT skills may facilitate the development of 21st-century skills in EFL classrooms. As a result of a collaborative problem-solving approach, DT significantly increased language learners' motivation, as well as their enjoyment and confidence about their communication and thinking skills. Furthermore, it provided students with greater autonomy and resulted in learning outcomes that were derived from the problem-solving process. As a result of repeated iterations of divergent and convergent thinking, students were able to practice and develop their ability to think about problems and discuss them in a team environment. A later study by Sotlikova (2023) found DT to be a powerful learning approach for transforming

ELT classrooms into hubs of innovation and problem-solving in an effort to boost English learners' motivation and empowerment, particularly within the field of CALL.

Despite the recent literature findings highlighting how DT could shape students' motivation, self-efficacy and engagement, particularly in STEM education, its role in language learning and language learners' psychological factors, such as their L2 motivational self-system and L2 grit, remains unclear. For clarification, Crites and Rye (2020) reported in their study that fostering DT in language classrooms can enhance language learners' creativity, collaboration and teamwork. In Roy's (2017) task-based language learning class, students learn language through Moodle, with a particular focus on improving their writing skills. The teacher indirectly guides students to incorporate DT into their writing, content organization, styling, spelling and grammar. It has been reported that through DT, language learners were able to learn not only DT fundamentals but also how to create complex technical documents in English with a high level of confidence. In another survey, Cleminson and Cowie (2021) applied DT in their language classrooms and reported that it enhanced language learners' writing skills, communication and creative collaboration in their language learning activities.

The literature review revealed that DT, which has been extensively applied in a variety of fields, particularly in STEM education, has positive effects on students' motivation and self-efficacy. However, only a few studies in applied linguistics have applied it to English language classes, and most of those studies indicated that it improved language learners' teamwork (Cleminson & Cowie, 2021), creativity (Crites & Rye, 2020) and writing skills (Nazim & Mohammad, 2022). To date, the influence of DT on psycholinguistic factors in our field has not yet been explored. Examining it could bring us closer to understanding its impact on language learners' motivation and L2 grit in AILL.

According to complex dynamic systems theory, several psycholinguistic factors contribute in various ways to the development of second language competence and skills (Larsen-Freeman & Cameron, 2008), which are closely related to learners' cognitive resources (e.g. learning strategies), motivation to learn, personality characteristics and psychological dispositions, especially those emphasized by positive psychology in the field of foreign language acquisition. In accordance with existing research, these factors include L2 grit, creativity, motivation, self-regulation, language mindsets, basic psychological needs, emotional intelligence and engagement in language learning, all of which affect learners' well-being (Dörnyei & Ryan, 2015; Teimouri et al., 2020). Cultivating any skill, particularly one related to DT, follows this rule, as learning a new language or skill in a new context can be challenging—especially when using

innovative tools like artificial intelligence in novel ways (e.g., cultivating DEAILL). This process may influence language learners' complex dynamic systems in language learning, particularly affecting their L2 motivation and L2 grit.

Recent L2 studies report that language learners must rely on cognitive, affective, and behavioral factors to learn a language in any context (Rahimi, 2023; Rahimi & Mosalli, 2024; Smith et al., 2020). In this regard, scholars have emphasized the importance of motivating learners to sustain their efforts and interest in language learning (Dörnyei & Ryan, 2015; Rahimi & Sevilla-Pavón, 2024; You & Dörnyei, 2015). This motivation can be fostered by integrating new CALL tools, such as AI, LMOOCs and online language learning, while also cultivating new skills and approaches alongside language proficiency. For clarification, Zheng et al. (2018) found that online language learning boosted learners' motivation, with their instrumentality-promotion and cultural interest influencing their online self-regulation. Following them, Smith et al. (2020) reported that learners' L2MSS, such as their ideal L2 self, ought-to L2 self, and current L2 self, drove their efforts in online language learning. Moreover, Rahimi (2023) reported the positive influence of language MOOCs (LMOOCs) in shaping Iranian language learners' instrumentality-promotion to achieve their personal objects and instrumentality-prevention to pass their academic rubrics, which shaped their continued intention to learn language in LMOOCs. He also underscored the predictive power of digital self-authenticity in shaping language learners' attitudes to learn language in LMOOCs which had already been reported to shape language learners' efforts in digital language learning (Henry & Cliffordson, 2017), online (Smith et al., 2020), and then their attitudes toward emergency remote language learning in Iran (Rahimi & Mosalli, 2024). In the context of Chat-GPT assisted language learning, Rahimi and Mosalli (2025) reported that language learners' perceived their ChatGPT-supported environment as a more authentic language learning environment, as they could solve real life challenges and language learning problems through it and this was more in line with their current and future self images. It remains to be seen how the context of artificial intelligence language learning, with new skills like DTCALL that will be cultivated alongside language skills, can influence language learners' L2 motivation.

Recent studies have also highlighted the impact of L2 grit on maintaining a student's interest in learning a foreign language as well as their effort. Like motivation, grit is known to be a context-specific factor, meaning that it behaves and is predicted differently in various language learning contexts (Pawlak et al., 2024). In this regard, recent L2 studies reported its crucial role in various language learning environments. For

instance, Chen Hsieh and Lee (2021) reported that language learners were grittier in robot-assisted language learning than when completing digital storytelling tasks; however, it had positively correlated with language learners' outcomes, particularly with their perseverance of effort. In the context of Virtual Exchange, Rahimi and Sevilla-Pavón (2025b) reported that learners' consistency of interest as well as persistency of effort were significantly mediated by the correlations between their L2 self-identities and Virtual Exchange self-regulation, which influenced their ability to evaluate learning progress, seek help, and set additional goals in this context. Despite emerging studies indicating a positive correlation between language-domain-specific grit and language learners' performance, little is known regarding the factors influencing the intensity of L2 grit and its predictors. This gap has been pointed out in previous studies stressing the pertinence of exploring it in various language learning contexts (Paradowski & Jelińska, 2023; Pawlak et al., 2022, 2024; Rahimi & Sevilla-Pavón, 2024). Accordingly, to heed these recommendations and take one step forward, we are now exploring how DTCALL, as a new 21st-century digital skill, might shape language learners' L2 grit in AILL.

Taking into account the established relationships between L2 achievement and motivation (Dörnyei & Ryan, 2015), some researchers have examined motivation as a predictor variable in relation to L2 achievement, typically under different terms (e.g. Pawlak et al., 2022). As these two constructs are intuitively related, it is not surprising to see such a focus on research. As Dörnyei and Ryan (2015) stated while a student may possess exceptional abilities, he or she will not be able to achieve long-term goals without sufficient motivation, especially when adequate curricula and instruction are provided. Griffiths and Soruç (2020) also asserted that 'motivation is indeed a sine qua non!' (As cited in Pawlak et al., 2024, p. 6). As a result of motivation, effort is invested in achieving goals (Melendy, 2008). The presence of high levels of motivation is also accompanied by high levels of grit (Pawlak et al., 2024; Rahimi & Sevilla-Pavón, 2025a). Furthermore, the level of grit exhibited by L2 learners—both generally and in the context of L2 learning—may be linked to motivation (Pawlak et al., 2024). Feng and Papi (2020) found that L2 grit is a strong predictor of language learners' performance, particularly when motivation serves as a mediator. This prompted us to examine the mediating role of motivation in the relationship between DEAILL and language learners' L2 grit in the present study. In light of previous findings and gaps in the literature, the researchers intend to address the following questions and test the following conceptual model, as shown in Figure 1.

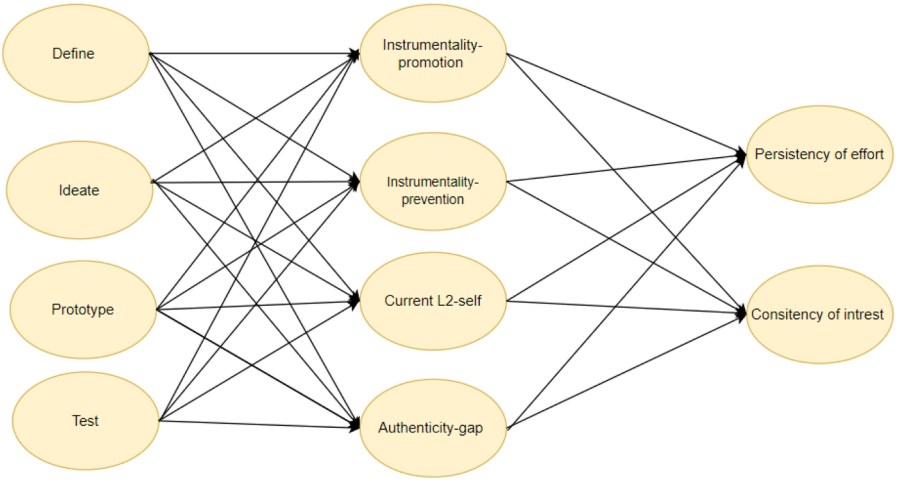

Figure 1. The study's hypothesized model.

- Q1. What are the factorial structures of Design Thinking skills in Artificial Intelligence Language Learning (DEAILL), Artificial Intelligence L2 motivational self-system (AIL2MSS), and L2 grit in the Spanish EFL context and Artificial Intelligence Language Learning?
- Q2. To what extent can DEAILL predict language learners' artificial intelligence L2 motivational self-system in AILL?
- Q3. To what extent can DEAILL predict language learners' L2 grit in AILL?
- Q4: To what extent could language learners' artificial intelligence L2 motivational self-system mediate or moderate the relationship between their DEAILL skills and L2 grit in AILL?
- Q5: How can DEAILL phases shape language learners' L2 motivational self-system and their L2 grit in artificial intelligence language learning?

3. Methodology

3.1. Study design

This study used a sequential mixed-method design, with the quantitative phase informing the qualitative phase (Creswell & Clark, 2011). The quantitative phase of the study involved the collection and analysis of numeric data through an online survey. Based on the survey results, the qualitative phase involved the analysis of essays written by students working in their local groups to provide deeper insights into the findings (Creswell & Clark, 2011).

Table 1. Participants' demographic information.

		N	%
Gender	Male	9	9.8
	Female	83	90.2
Age	18-22	79	83.7
	23-26	7	7.6
	26<	6	6.5
Language Learning Experience	Years		
	1-2	2	2.2
	3-5	72	78.3
	4<	18	14.1

3.2. Participants

The study included 92 participants who were enrolled in a 'History and Culture of English-speaking Countries' course taught by the second author at the Universitat de València in Spain. Demographically, there was a larger number of females (83), which was followed by a small number of males (22). 77 participants were between 18 and 22, seven were between 23 and 26, and six were over 26. Table 1 below summarizes participants' demographic information.

3.3. Research procedure

The study uses data from the VEMO-Hist" iteration within SOCIOEMOVE, a collaborative project among language learners from different cultures and languages who worked in groups of four to five students using different AI tools and apps throughout the fall semester of 2023. As participants were enrolled in a 'History and Culture of English-speaking countries' university course, all the data, fact and pieces of information they worked with to produce their artifacts were based on reality, as specified in the syllabus. First of all, the new countries participants were asked to create had to be based on features taken from the different geopolitical, social and legal systems of English-speaking countries. This means students were required to learn about those real-world system and countries beforehand. Secondly, the historical figures students had to conduct research about had all existed and had made significant contributions to the history and culture of English-speaking countries, thus aligning with the course syllabus. It is also worth mentioning that the AI tools used enabled participants to have live conversations with those historical figures, in the same way they would have done in real life.

The Project began with the 'Getting Started' phase, involving self-introduction videos, questions and intercultural ice-breaking games in the Google Classroom forum with the support of AI tools (e.g. Synthesia for creating videos using AI avatars, ChatGPT to generate questions or prompts for discussions in Google Classroom and for

receiving real-time answers and automatic feedback, and Grammarly to get feedback on grammar and style). These tools were used for language assistance, information search and content organization. This was followed by the DEAILL 'Define' phase, where participants were encouraged to identify the areas that required additional assistance in terms of cultural, historical and linguistic learning. Next, in the DEAILL 'Ideate' phase, the HelloHistory AI app was used by students in order to interview a historical figure from their target culture. This app enabled them to have conversations with their chosen historical figures in the second person singular as if they were right in front of them to obtain information they could use when brainstorming, highlighting the key concepts and feedback provided by AI and discussing possible educational, economic, political, and and legal systems, as well as tourism industries, cultural practices, official languages, etc. in their ideal new countries. Following this, the DEAILL 'Prototype' phase invited participants to summarized the features of their ideal new country, highlighting what made it unique and how it met perfect living standards, based on the previously AI-supplied feedback. Once they had summarized their ideal country's main features, it was time to bring this new country's concept to life in an audiovisual format with the support of different AI tools (e.g. Text-to-Image AI tools such as DALL-E and Midjourney to create images representing landscapes, architecture and public spaces of the ideal country; 3D Modeling AI tools such as Blender with AI-based plugins to render immersive, realistic environments and cityscapes; Narrative AI such as ChatGPT when scripting their narratives, including dialogues or voiceovers while providing structure to the digital story; Voice Synthesis AI such as Descript or ElevenLabs for voiceovers in multiple languages and accents to enhance inclusivity; and video editing and animation AI tools such as Adobe Premiere Pro and the Runway AI-based video editor to animate characters and scenes, and to blend visuals with audio for a polished final artifact). Thus, after collecting information about the historical figure and the ideal features for their country, they wrote scripts for their final audiovisual artifacts: digital stories. AI-assisted voiceover was used in students' digital stories (fictitious documentaries about a historical figure or new nation, or commercials encouraging people to move to the new nation). At this stage, learners collected AI-generated feedback and suggestions for script improvement from AI writing assistants. Finally, the DEAILL 'Test' phase involved evaluating their artifacts based on the feedback received from their target context, including peers and teachers on Google Classroom: students exchanged comments, suggestions for improvement and constructive criticism and, once the final artifacts were generated, they voted to choose the best

ones as part of the 'Wrapping-up and Celebrating' Phase, where the best artifacts received prizes in an award-giving ceremony held at the end of the semester.

3.4. Instrumentation

The research instrumentation for this study consisted of three components, which are attached as Appendix A. The first block gathered language learners' demographic information regarding their gender, age, language learning experience and informed consent. The second block explored language learners' motivation. It was adated from Rahimi (2023) and Smith et al. (2020) and included the following four items: instrumentality-promotion, instrumentality-prevention, authenticity gap and current L2 self. For clarification, we changed the term LMOOC to AI in 'Learning English in this LMOOC is more interesting than a face-to-face language class'. Regarding instrumentality-promotion, the item was revised from 'Learning English in this LMOOC is important to me because other people will respect me more if I have a knowledge of English' to 'Learning English with Artificial Intelligence is essential for me because acquiring the language will change my life in the future'. For instrumentality-prevention, the item was revised from 'I have to learn English in this LMOOC to pass my final semester' to 'I have to learn English with artificial intelligence to pass my final semester'. Regarding current L2 self, this item was revised from 'From now on, I see myself as someone who is good at speaking/using English in online language learning' to 'Now, I see myself as someone who can use artificial intelligence to learn language in many different situations'.

The third part of our instrument is DEAILL, which was adapted from Tsai and Wang (2020) and Stanford D. School (2010) and comprises four latent variables: Define, Ideate, Prototype and Test. For instance, we changed 'I usually generate solutions *via* brainstorming' to 'I brainstorm to build on the constructive ideas and information generated by artificial intelligence'. 'I usually have a clear idea about the problem that I am facing' was revised too: 'I usually try to clearly identify my current and future language needs that can be met by using artificial intelligence'. Finally, the last part of the study scale, which tapped into language learners' L2 grit, included items adapted from Teimouri et al. (2022) and evaluated language learners' persistence of effort and consistency in interest in AILL. The items in all the sections except from the first one ranged from 1 (strongly disagree) to 5 (strongly agree). The descriptive statistics provided in Appendix A pertain to the DEAILL, language learners' motivation and L2 grit latent variables. As displayed on the table, the average mean scores of these variables were above 3, indicating that most

participants chose the top Likert options. In spite of this, it is important to note PLS-SEM does not have a high sensitivity to normality distribution. Nevertheless, the data was normally distributed for both skewness and kurtosis, which was between −1 and 1.

In addition, the second author of this paper, who designed and facilitated the study project, and was also in charge of collecting the data, developed a series of essay prompts based on the DEAILL framework and its phases. The Delphi methodolgy was applied in order to assess content and face validity of the questions students were asked to address in their essays. Accordingly, three expert researchers from the fields of CALL, Psycholinguistics and STEM anonymously evaluated and validated the questions, ensuring *via* two rounds of review that they were aligned with the construct and thus the questions they comprehensively covered the key aspects of the research topic while aligning with the study's goals. The questions were first piloted with a small group of students before being administered to 81 participants who wrote essays. This way, the clarity of the questions was ensured.

The essay questions were addressed in the form of essays wrote by 81 participants The essay questions were used to gather detailed insights into how DEAILL can shape language learners' L2 motivational self-systems and L2 grit across each phase of the DEAILL framework. After the second author devised the initial essay prompts, the first author and three CALL experts reviewed them to ensure face and content validity. They evaluated the prompts for conceptual relevance and their alignment with CALL, AILL, Psycholinguistics, and Applied Linguistics. Furthermore, the protocol was evaluated in a pilot study with three participants with similar profiles to those of the main study, who reported having no difficulty understanding the different essay prompts.

3.5. Data analysis

The Statistical Package for Social Sciences (SPSS) and PLS-SEM (4.0) were implemented to achieve the study objective. We selected the PLS-SEM as it is a superior data analysis approach that follows composite-based structural equation modeling, has a higher level of statistical power, and shows a more consistent convergence pattern than the CB-SEM with small sample sizes (Hair et al., 2019). Moreover, its predictive nature makes it particularly suited for theory development (Hair et al., 2023). PLS-SEM enables the specification of latent variable interactions, including direct, indirect, and serial mediation (Hair et al., 2023). Lastly, Sarstedt et al. (2020) asserted that this approach is superior to CB-SEM in terms of estimating mediation. Accordingly, first of all, we investigated the factorial validity of our variables through exploratory factor analysis

(EFA), confirmatory factor analysis (CFA), and the PLS-SEM formative phase. Next, in the reflective phase, we examined correlations, directions, and serial mediations among variables, while also evaluating the fit of the study model.

Braun and Clarke's (2006) guidelines were adapted for reflective thematic context analysis using MAXQDA (20.0). After collecting the essays, they were read several times by the researchers so that they could become familiar with the data. Following that, a 'theoretical' thematic or deductive content analysis was performed. This way, quantitative results informed further qualitative analyses aimed at identifying language learners' L2 self-identities and L2 grit in the different DEAILL phases. Based on each DEAIL phase, the qualitative data were analyzed accordingly by the second author, who had also designed the project tasks, facilitated the exchange and collected the data. In order to ensure inter-coder reliability, an expert qualitative analyst also reviewed and analyzed the qualitative portion of the study (Creswell, 2014).

4. Results

4.1. Exploratory factor analysis, confirmatory factor analysis (CFA), and measurement model

Following the collection of data, the scale items were dimensionally assessed using EFA through SPSS. This is accomplished by measuring Kaiser-Meyer-Olkin (KMO) and Bartlett's sphericity test. A KMO value greater than 0.50 indicates that the data set is suitable for factor analysis. The Bartlett value obtained was below 0.85, thus rejecting the null hypothesis at the 0.05 significance level. After that, principal component analysis and varimax steep rotation were applied as factorization techniques. Factor analysis can be used to determine whether items on a scale are distributed among fewer factors or not. Accordingly, factor loads less than 0.5 and initial eigenvalues less than 1 should be excluded from the questionnaire. In Table 2, all four of DEAILL's factor loadings exceeded 0.5, and their initial Eigenvalues exceeded 1. Furthermore, the factor loads of the items on the scale explained at least 44% of the variance, which is considered sufficient in a humanistic field.

Regarding the language learners' AIL2MSS scale, the KMO obtained 0.86, and the Bartler test obtained 0.66. Using principal component analysis and Varimax steep rotation as factorization techniques, all the factor loads were greater than 0.5, and all of the initial eigenvalues of four constructs were greater than 1, as indicated in Table 3. Furthermore, they are able to predict at least 44% of the variance in motivation of participants in AILL.

Table 2. The EFA analysis of the DEAILL.

	Factors				
Factors	DE	PR	ID	TE	Extraction
DE3	**0.842**	0.240	−0.011	−0.189	0.803
DE2	**0.825**	0.255	0.102	−0.124	0.771
DE1	**0.841**	0.117	0.019	−0.075	0.727
PR1	0.115	**0.855**	0.208	−0.061	0.791
PR2	0.153	**0.850**	0.276	0.082	0.829
PR3	0.147	**0.842**	0.224	0.193	0.818
ID1	0.173	0.248	**0.849**	−0.012	0.813
ID2	0.216	0.213	**0.856**	0.031	0.825
ID3	0.140	0.305	**0.853**	−0.099	0.849
TE1	0.263	0.037	0.421	**0.708**	0.717
TE2	0.201	0.023	0.294	**0.749**	0.669
TE3	0.145	0.011	0.227	**0.779**	0.666
Initial Eigenvalues	5.65	2.25	1.31	1.01	
Percentage variance explained	43.17	17.34	10.07	7.78	
Percentage variance Cumulative	44.17	60.46	70.54	78.32	

Table 3. The EFA analysis of the AIL2MSS.

	Factors				
Factors	IPR	IPO	CU	AUT	Extraction
PR1	**0.865**	0.198	0.256	−0.013	0.813
PR2	**0.873**	0.235	0.112	0.242	0.833
PR3	**0.882**	0.323	0.096	0.181	0.820
IPO1	0.040	**0.654**	0.302	0.478	0.582
IPO2	0.046	**0.724**	0.229	0.476	0.632
IPO3	−0.087	**0.736**	0.315	0.423	0.629
CU1	0.105	0.252	**0.764**	0.319	0.696
CU2	0.183	0.124	**0.826**	0.165	0.743
CU3	0.236	0.364	**0.834**	0.170	0.780
AUT1	0.265	0.163	0.086	**0.786**	0.696
AUT2	0.239	0.352	0.248	**0.807**	0.771
AUT3	0.161	0.214	0.329	**0.732**	0.669
Initial Eigenvalues	5.52	1.96	1.82	1.16	
Percentage variance explained	46	25.12	15.61	9.6	
Percentage variance Cumulative	46%	61%	52%	71%	

In relation to the EFA analysis for L2 grit scale, the KMO obtained 0.70, and the Bartler test obtained 0.15. As a result of applying principal component analysis and varimax steep rotation as factorization techniques, all the factor loads were greater than 0.5, and the initial eigenvalues values for CI and PE were greater than 1, as indicated in Table 4. Furthermore, the factor loads of the items on the L2 grit scale explained at least 47% of the variance, which is considered sufficient in a humanistic field.

To determine the DEAILL's factorial structure, we ran CFA in the form of CB-SEM before moving to the PLS-SEM's measurement model or reflective phase. The CFA result showed an adequate model fit to the data ($\chi 2/df$ = 1.41; the Root Mean Squared Error of Approximation (RMSEA) = 0.06; the Normed Fit Index (NFI) = 0.91; the Comparative

Table 4. The EFA analysis of the L2 grit.

	Factors		
Factors	CI	PE	Extraction
CI1	**0.823**	0.121	0.693
CI2	**0.835**	0.242	0.756
CI3	**0.840**	0.044	0.707
PE1	0.247	**0.730**	0.594
PE2	0.086	**0.889**	0.797
PE3	0.069	**0.888**	0.793
Initial Eigenvalues	2.82	1.51	
Percentage variance explained	47.12	25.21	
Percentage variance Cumulative	47.12	72.33	

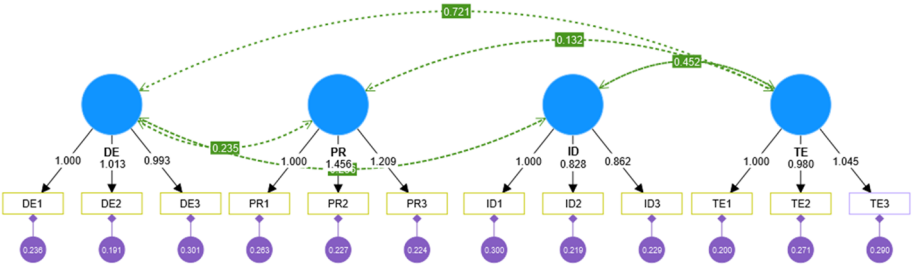

Figure 2. The factor loadings of DEAILL.

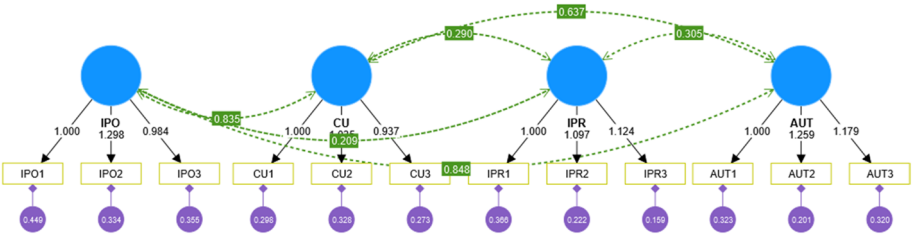

Figure 3. The factor loadings of AIL2MSS.

Fit Index (CFI) = 0.97; and the Tucker Lewis Index (TLI) = 0.962) as presented in Figures 2. The factor loadings of the four indicators ranged from 0.99 to 1.22, which is acceptable and shown in Figure 2. Similarly, the CFA results of the OL2MS, ($\chi 2/df$ = 1.06; RMSEA = 0.01; NFI = 0.92; GFI = 0.91; TLI = 0.99), and for L2 grit ($\chi 2/df$ = 1.47; RMSEA = 0.07; NFI = 0.94; GFI = 0.98; TLI = 0.96), showed an adequate model fit with high factor loadings, which are shown in Figure 3 and 4, respectively.

Regarding the PLS-SEM reflective model, which evaluates variable factor loadings, reliability, convergent validity, and discriminant validity, we used Cronbach's alpha and composite reliability to assess reliability. Hair et al. (2023) recommend that these values be above 0.70. For convergent validity, the average variance extracted (AVE) should exceed 0.5. As shown in Table 5, the latent variables were both reliable and valid.

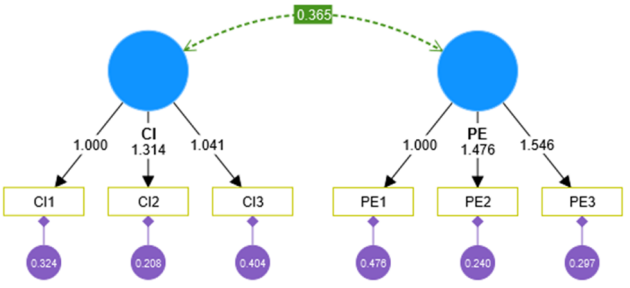

Figure 4. The factor loadings of L2 grit.

Table 5. Validity and reliability of the study latent variables.

Constructs	Cronbach's alpha	Composite reliability (rho_a)	Composite reliability (rho_c)	Average variance extracted (AVE)
AUT	0.837	0.839	0.902	0.755
CI	0.800	0.808	0.882	0.714
CU	0.847	0.848	0.907	0.766
DE	0.886	0.891	0.929	0.814
ID	0.901	0.906	0.938	0.835
IPO	0.734	0.735	0.849	0.653
IPR	0.917	0.917	0.948	0.858
PE	0.805	0.805	0.885	0.720
PR	0.908	0.909	0.942	0.845
TE	0.857	0.863	0.913	0.777

Table 6. Heterotrait-monotrait ratio (HTMT) - Matrix.

	AUT	CI	CU	DE	ID	IPO	IPR	PE	PR	TE
AUT										
CI	0.731									
CU	0.639	0.684								
DE	0.465	0.356	0.435							
ID	0.659	0.573	0.684	0.350						
IPO	0.863	0.738	0.830	0.324	0.731					
IPR	0.325	0.372	0.304	0.364	0.274	0.232				
PE	0.573	0.387	0.417	0.362	0.242	0.604	0.225			
PR	0.660	0.642	0.553	0.571	0.359	0.606	0.230	0.531		
TE	0.756	0.530	0.716	0.350	0.752	0.759	0.328	0.312	0.499	

Discriminant validity was assessed using the Heterotrait-Monotrait ratio (HTMT) proposed by Henseler et al. (2014), which measures whether a construct is empirically distinct from other constructs in the model. According to Henseler et al. (2014), the mean HTMT value should not exceed 0.85 or 0.90. The results, shown in Table 6, support this criterion, thus answering the first research question of the study.

4.2. The structural model

The evaluation of the formative model includes several metrics, such as multicolinearity, path coefficient effect size, coefficient of determination (R^2), Stone-Geisser's Q^2 value, and Standardized Root Mean Square Residual (SRMR).

As the first metric, the collinearity was measured through the variance inflation factor (VIF). According to Hair et al. (2023), the VIF indices for all of the endogenous variables should be below 5. To answer the second research question, we conducted an analysis of the direct path coefficients in the structural model to examine the correlation between DEAILL and learners' AIL2MSS in AILL. According to Table 7, The paths between DE->IPR († = 0.269;), ID->IPO, († = 0.342), ID->CU († = 0.315), ID->AUT († = 0.228), PR->IPO († = 0.308), PR->CU († = 0.221), PR->AUT († = 0.324), TE->IPO († = 0.273), TE->CU(† = 0.285), and TE->AUT(† = 0.331) were supported, while the remaining paths had either a t-value below 1.96 or a p-value above 0.05, indicating they were not significant.

To address the study's third research question, a direct analysis with 5000 subsamples was conducted to explore the direct effect of DEAILL on language learners' L2 grit. As shown in Table 8, all DEAILL skills were found to directly predict participants' L2 grit, except for the Define skill. Specifically, Ideate (DE -> CI; β = 0.215), Prototype (PR -> CI; β = 0.208), and Test (TE -> CI; β = 0.247) positively predicted learners' CI.

Furthermore, the bootstrap indirect analysis revealed that only the authenticity gap mediated the relationship between learners' test and CI (TE -> AUT -> CI; β = 0.106), as well as between prototype and CI (PR

Table 7. Result of the structural model.

Path exogenous --------> endogenous	β	t-value	p-value	VIF
DE->IPO	−0.083	0.902	0.367	1.40
DE->IPR	0.269	2.351	0.019	1.40
DE->CU	0.075	0.977	0.329	1.40
DE->AUT	0.064	0.719	0.472	1.40
ID->IPO	0.342	3.080	0.002	1.82
ID->IPR	0.049	0.339	0.734	1.82
ID->CU	0.315	2.830	0.005	1.82
ID->AUT	0.228	2.398	0.017	1.82
PR->IPO	0.308	3.090	0.000	1.53
PR->IPR	−0.025	0.211	0.833	1.53
PR->CU	0.221	2.584	0.010	1.53
PR->AUT	0.324	3.595	0.000	1.53
TE->IPO	0.273	2.766	0.006	1.97
TE->IPR	0.189	1.239	0.215	1.97
TE->CU	0.285	2.486	0.013	1.97
TE->AUT	0.331	3.493	0.000	1.97

Table 8. Result of the bootstrap direct analysis.

Path	β	2.5%	97.5%	t-value	p-value
DE->CI	0.058	−0.045	0.157	1.123	0.262
DE->PE	0.012	−0.098	0.126	0.217	0.828
ID->CI	0.215	0.089	0.346	3.280	0.001
ID->PE	0.158	0.036	0.270	2.658	0.008
PR > CI	0.208	0.098	0.341	3.356	0.001
PR->PE	0.171	0.069	0.302	2.872	0.004
TE->CI	0.247	0.119	0.364	3.941	0.000
TE->PE	0.176	0.068	0.295	3.025	0.002

-> AUT -> CI; β = 0.104), as shown in Table 9. The remaining variables were non-significant and could not act as mediators between DEAILL's correlation with learners' L2 grit, thus functioning as moderators instead. This allowed us to answer the study's final research question.

The coefficient of determination was used to estimate the contributions of exogenous and mediator variables to endogenous variables, thus measuring the model's predictive power. According to Hair et al. (2023), prediction power is considered weak, moderate, and strong at values of 0.19, 0.37, and 0.677, respectively, all of which were achieved at sufficient levels in our model. Additionally, the Stone-Geisser (Q^2) was calculated to assess predictive accuracy, with an acceptable relevance level defined as a value above 0, which was met in this study. The SRMR was below 0.08, indicating a good model fit (Hair et al., 2023), as shown in Table 10. Figure 5 presents the study model with the corresponding relative values.

4.3. Essays

18 final essays, written by students working in local groups of 4 or 5 once the different project tasks had been completed, were analyzed to

Table 9. Indirect analysis.

Path	β	2.5%	97.5%	t-value	p-value
PR-> IPR ->PE	−0.002	−0.036	0.038	0.090	0.928
TE->IPR->CI	0.024	−0.024	0.085	0.883	0.377
TE->AUT->CI	0.106	0.036	0.210	2.377	0.018
DE->AUT->PE	0.018	−0.040	0.074	0.650	0.516
TE->IPR > PE	0.011	−0.046	0.088	0.364	0.716
ID->CU->CI	0.076	0.007	0.172	1.791	0.073
PR->AUT->CI	0.104	0.028	0.214	2.145	0.032
DE->CU->PE	0.001	−0.026	0.034	0.038	0.969
TE->CU->CI	0.069	0.002	0.159	1.712	0.087
ID->IPO->CI	0.060	−0.016	0.156	1.370	0.171
PR->IPR->CI	−0.003	−0.040	0.039	0.172	0.863
DE->AUT->CI	0.020	−0.037	0.084	0.676	0.499
PR->AUT->PE	0.090	0.000	0.222	1.529	0.126
ID->IPR->PE	0.003	−0.041	0.048	0.142	0.887
TE->CU->PE	0.002	−0.082	0.083	0.052	0.958
TE->IPO->PE	0.072	−0.021	0.182	1.420	0.156
PR->CU->PE	0.002	−0.055	0.073	0.051	0.959
PR->CU > CI	0.054	0.004	0.139	1.496	0.135
ID->AUT->CI	0.073	0.008	0.161	1.857	0.063
ID->IPO->PE	0.090	−0.020	0.234	1.363	0.173
DE->IPO->PE	−0.022	−0.092	0.033	0.707	0.480
PR->IPO->CI	0.054	−0.013	0.145	1.318	0.187
TE->IPO->CI	0.047	−0.016	0.120	1.399	0.162
ID-> IPR -> CI	0.006	−0.034	0.063	0.271	0.787
DE ->IPR->PE	0.016	−0.051	0.102	0.438	0.661
ID->AUT->PE	0.063	−0.003	0.160	1.434	0.152
DE->IPO->CI	−0.014	−0.062	0.021	0.709	0.478
TE->AUT->PE	0.091	0.000	0.213	1.687	0.092
DE-> CU->CI	0.018	−0.020	0.065	0.865	0.387
PR->IPO->PE	0.081	−0.017	0.219	1.332	0.183
ID->CU->PE	0.002	−0.098	0.081	0.053	0.958
DE->IPR ->CI	0.034	−0.013	0.100	1.148	0.251

validate and supplement the questionnaire quantitative results. In this regard, in the Define phase, the students were primarily focusing on their study project and their course overall, showing that initially they were mainly concerned with succeeding in passing the course and obtaining a satisfying final project outcome. Their responses showed that in this phase they displayed high instrumentality-prevention.

Group 7

At first we thought we had to keep talking about the former project (the interview to a historical figure).

Group 2

We only used AI for the activity we had to do about asking a famous history character

Table 10. Result of the structural model.

Variables	R^2	Q^2
AUT	0.557	0.405
CI	0.469	0.319
CU	0.498	0.363
IPO	0.501	0.302
IPR	0.151	0.113
PE	0.267	0.149
SRMR = 0.05		

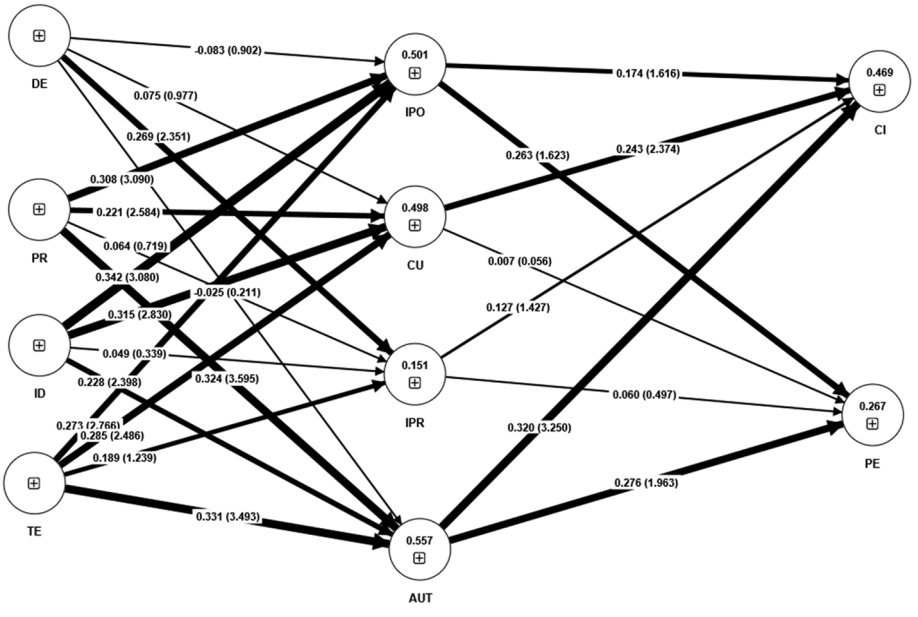

Figure 5. The structural model with relative values.

Group 5:

> The use of AI, at least in our group, is limited to searching for the best sites to get information, in fact it can give you the best ways to learn about every topic, such as books, cultural sites and films. In the solving process, we use it as a help to find the ways that people before us solved the problem first; then we can elaborate our own way to solve it, inventing a whole new way.

Regarding the second phase of DEAILL, the findings show that learners' reported that AI had given them even more information than the amount required, which had not been provided as part of their university courses, showing the authenticity gap, and now they had a lot of information about what their current needs were (current L2 self). Moreover, sharing some information about one of the participants' historical figures led her to display a high level of L2 grit.

Group 9:

> We use AI in almost all of our works, but we do not abuse it, using it to answer all questions and not even looking at the answers. We use it to ensure that our answers are correct and to compare many answers and many words that can be used in our work. We usually try to integrate many new words given by it, to improve our vocabulary in the essays or worksheets and in our daily English activities.

Group 10

> "We tend to ask AI programs for help. Also, as part of this lecture, we were made to work with the HelloHistory AI App, a program to interview historically important persons. In general, it caught our interest, as it gave us the opportunity to work with a program different than the AI program usually used"…… AI was helpful to acquire useful information for our topics, gaining knowledge a lot faster, and better that we might not know. Additionally, AI could serve as a tool to relieve the work of the teachers, as students are confronted with more independent learning as a consequence of more AI being used.

For the following participants, the findings show that in the Ideate phase the AI shaped language learners' current L2-self, instrumentality promotion, L2 grit and authenticity gap:

Group 9:

> We envision streamlined processes in future projects, coupled with the autonomy to submit individual assignments. This autonomy aligns with our values, believing it would positively contribute to project efficiency and quality, allowing for a more balanced and productive experience.

For the following participants, the findings show that in the Ideate phase AI shaped language learners' current L2 self and authenticity gap:

Group 10

We used the HelloHistory AI app, which enables us to talk with a historical figure of our choosing. Using this app really improved our understanding of that person because it made the experience and the persona more realistic. It talked about the person's thoughts and feelings, and provided more context of the situations that this person was in. From translating a text to checking the grammar of an essay or conversing with a historical figure, AI can be a valuable tool for students as long as it is used properly and with responsibility.

Regarding prototype phase, the findings showed it had significantly shaped participants L2 grit, authenticity-gap and current L2 self, as illustrated by the following observations made by different groups of students in their essays:

Group 3

We try to solve any problem by ourselves but it is true that when facing an issue that we cannot find an answer to we ask an AI for advice or alternatives. For example, when we had doubts about a tough question and even comparing knowledge between each other we could not answer it we asked GPT to tell us more about the topic so then we could learn more about it and solve the question. Because of all these reasons, using AI moderately does not negatively affect the student and does not harm them.

Group 2

We loved the experience with the HelloHistory AI App because we were excited to talk to a historical figure like Shakespeare. We love their work and the repercussions that he had made around the ages. I would recommend this experience to everyone

Group 3

It was such an amazing experience using the HelloHistory App. We had no idea that this kind of application existed and we had such a great time with it. Not only using it for our project but also out of curiosity

Group 9

Our experience using the Hello History application can be described as both fun and interesting. It was especially curious to see how a simple AI tool could perfectly simulate the possible answers that a historical figure -Rosa Parks, in our case- would give according to her behavior, personality and history.

Group 7

We began to work on our own video to encourage people to move to Klepertopia, for which we decided to use AI generated voices and some videos we found useful to depict the topic that was talked about

Finally, in the testing phase, the results demonstrated a significant impact on learners' authenticity gap, current L2 self, and L2 grit. For example, participants reported that:

Group 7

> We only used AI for the activity we had to interview a famous historical character. We used "HelloHistory" and we found this app very interesting because it is kind of unusual to talk with someone who is dead or difficult to talk to… AI generated voices and some videos we found useful to depict the topic that was talked about "As a group, we think we worked very well together and we have been able to face the problems and to express ourselves easily. Finally, we hope that future generations make correct use of AI, since it could successfully improve our current society"

Group 1:

> The role of artificial intelligence in our current era is relevant because of the many uses it has; to search information you don't know, complete assignments, develop your skills, enrich your interests and many others. In our case, thanks to the university procedure that we are adapting to, we have begun to learn about its many benefits, thus, we have introduced it in our learning process. However, we are very aware of the dangers it has, and one of the most concerning ones is how it can easily be misused and believed to have the same role as a teacher. What we mean with this, is that some people, when consulting the doubts, they have about a topic they are learning about, resort to different forms of AI instead of a teacher, who is a proper professional about the subject.

Group 4

> Throughout our experience, we must highlight how helpful Artificial Intelligence has been with the process of creating a new country as well as for meeting new people from a different country. These tools have facilitated interactions between us, Spanish students and the Cyprus students, enabling contacts despite the difficulties of synchronizing our schedules due the time difference. Despite the physical separation, it has streamlined communication and project management enabling a more efficient collaboration.

Group 9

> We have been able to better understand the content of the lessons thanks to the AI as some of the content that we have learnt we have been able to discuss it and apply it with our mates from Ireland. For example, we have practiced our contents in together at the time we needed to explain our scripts and give each other back up.

For the following participants, the findings showed that in this phase AI shaped their L2 grit, particularly learning with it high interest

Group 14

All in all, despite many people being against using AI tools for learning purposes, we think that we should not see them as a threat, but as an opportunity, as long as we use them with responsibility. We are looking forward to working with these tools again in the near future.

Group 3

We used the AI tool of "HelloHistory" to help us write a script for our Emu War project for our VE. The tool was easy to use with an easy to understand user interface (UI) and the result were exactly what we were looking for. Tools such as "HelloHistory" can be a great asset in education as they are easy to use and accessible.

5. Discussion

5.1. The factorial structure of DEAILL, AIL2MSS and L2 grit

The measurement model validated our new theoretical framework and instrument, Design Thinking Skills in Artificial Intelligence Language Learning (DEAILL), which comprises four components in this field and within the Spanish context, as shown in Figure 6. In response to the recommendations of Rahimi (2024), Rahimi and Sevilla-Pavón (2025b), and Rahimi and Mosalli (2024) for CALL researchers and educators to integrate 21st-century digital skills alongside language teaching expertise, we are currently validating DEAILL as one such skill in the field. This marks a significant advancement, also driven by the call from prominent figures in CALL to progress with technological advancements (Colpaert, 2020; Gimeno-Sanz, 2015).

Furthermore, following previous studies which validated the new AIL2MSS and L2 grit in online language learning (Rahimi & Sevilla-Pavón, 2025a; Zheng et al., 2018), as well as LMOOC (Rahimi, 2023) in Iranian and Chinese contexts, we validated these factors in Spanish and AILL contexts as well. The AIL2MSS factors, such as authenticity-gap and current L2 self, had previously been validated in Sweden (Henry & Cliffordson, 2017), Iran (Rahimi, 2023), Indonesia (Lamb & Arisandy, 2019), and China (Smith et al., 2020). However, they had not previously been validated in a Spanish setting, which is now addressed by this study. In addition, by validating L2 grit, we confirm the findings of previous studies regarding the validity of specific L2 grit factors (Alamer, 2021; Li & Yang, 2023; Teimouri et al., 2022). The validation of these frameworks and variables is crucial for the study's context, as future research in the Spanish context, particularly in AILL, can build upon this framework without the need for additional validity assessments.

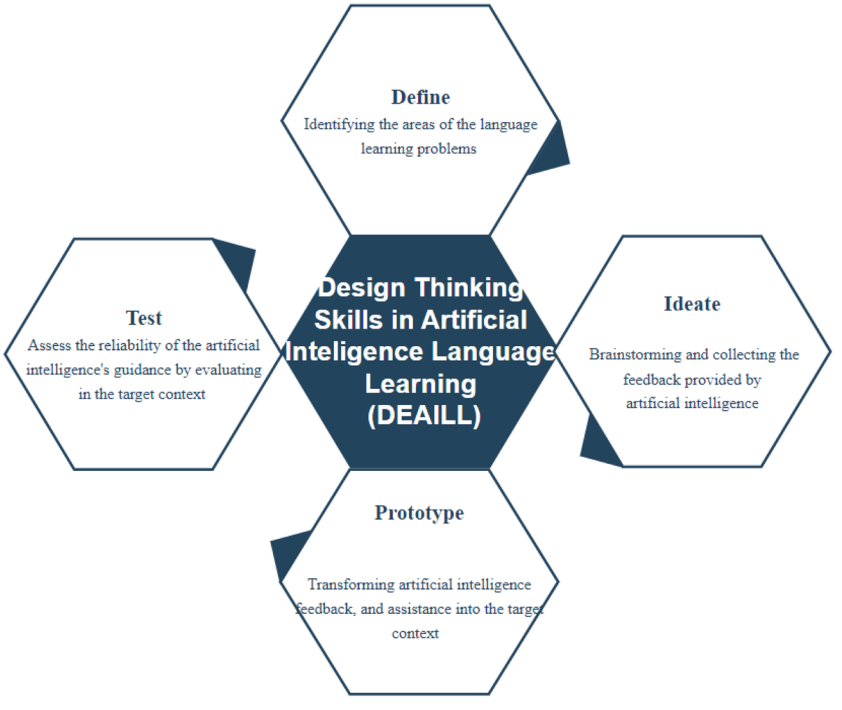

Figure 6. The DEAILL components.

5.2. Relationship between DEAILL and language learners' artificial intelligence L2 motivational self-system

The result of the structural model showed that language learners Defineskill only predicted their instrumentality-prevention ($\beta = 0.269$), suggesting that AI was used to help learners identify weak areas and fulfill their language learning obligations, such as completing university courses and obtaining a passing grade in such courses. While recent concerns have been raised about students using AI for completing academic tasks or even cheating (Hays et al., 2024), we view this as a positive development. It enables learners to pinpoint areas they may have overlooked, helping them overcome these difficulties, which, in turn, motivates them to learn languages through AI—requiring them to identify these issues beforehand. This aligns with Rahimi and Sevilla-Pavón's findings (2025). Define, however, failed to predict learners' current L2 self, likely due to both the phase's characteristics and learners' existing L2 self-guides. Stanford D. School (2010) suggests that the Define phase encourages design thinkers to confront challenges in their environment, and since our participants were focused on completing their university course requirements, this might explain why Define correlated only with instrumentality-prevention (IPR) and not the current L2 self. This was further supported by our qualitative results, where learners indicated that

the Define phase was primarily focused on completing tasks for the project and obtaining passing grades.

Furthermore, learners employed divergent thinking during the Define phase (Lin & Chang, 2024), gathering a vast amount of information related to the problem at hand. Rahimi and Sevilla-Pavón (2025a) also highlighted that learners' current L2 selves evolve in AILL as they progress with their projects, facing challenges along the way. The more learners engage with learning problems and shift from divergent to convergent thinking, the more they evaluate their progress. Thus, as learners moved from divergent (Define) to convergent (Prototype) thinking during DEAILL, their motivation shifted from prevention to promotion and present-oriented.

Moreover, learners' Ideate skills significantly predicted their instrumentality-promotion ($\beta = 0.342$), current L2 self ($\beta = 0.315$), and authenticity gap ($\beta = 0.228$). This indicates that AI feedback played a significant role in helping learners shape their ideal future selves, allowing them to achieve personal language learning goals and enhance their present self-description about their AI-supported language learning process. The promotional aspects of language learning in the Spanish context, such as learning a language for work opportunities or cultural competence (Sevilla-Pavón & Nicolaou, 2020), may have contributed to this.

Additionally, the personalized nature of AI, providing feedback based on individual needs, helped learners address specific learning difficulties, such as issues related to historical figures or ideal countries. This customization made the learning experience more authentic than previous language learning contexts, increasing learner satisfaction and motivation by enabling them to meet both their general and specific needs (Rahimi & Mosalli, 2025). In the Ideate phase, AI feedback significantly correlated with the authenticity gap, meaning that learners perceived AI-generated feedback and assistance as more authentic compared to prior learning experiences. This aligns with Rahimi et al. (2025), who found that Iranian EFL learners considered ChatGPT-assisted learning more authentic than their previous language learning contexts, particularly when it came to seeking assistance. This finding further enriches the CALL literature, as previous studies have demonstrated that Ideate skills significantly contribute to STEM learners' motivation (Tsai & Wang, 2020) and self-efficacy (Tsai et al., 2023). Additionally, our qualitative results support these findings, with students reporting that the AI-provided information and feedback enhanced their language learning experience and boosted their motivation to use such tools, particularly in terms of instrumentality-promotion and the authenticity gap. Students also noted in their essays that the AI's

insights were captivating, as they had not previously considered the information provided, suggesting that they reflected on their existing knowledge of culture and history during this phase (current L2 self).

Furthermore, our results indicated that Prototype skills significantly predicted language learners' instrumentality-promotion (β = 0.308), current L2 self (β = 0.221), and authenticity gap (β = 0.324). This suggests that applying AI-generated feedback and information in their language learning context enhanced learners' perception of their future language learning goals while increasing their satisfaction with their current language development. Recent CALL studies have also highlighted the significant impact of AI-provided feedback on language learners, particularly in terms of customized learning (An et al., 2023), motivation (Jeon, 2022), self-regulation (Rahimi et al., 2025), and behavioral intentions to use AI-driven tools (An et al., 2023).

Our study contributes to the literature by highlighting the practical implications of AI-generated feedback in shaping language learners' DEAILL, which in turn enhanced their instrumentality-promotion (IPO) and current L2 self. Additionally, we observed a high level of shared variance between prototype skills and digital self-authenticity, which may be attributed to the "engaging, meaningful, and personalized" digital learning opportunities identified by Rahimi et al. (2025) as key factors in bridging the authenticity gap between AILL and traditional language learning contexts.

The last factor of DEAILL significantly predicted learners' instrumentality-promotion (β = 0.273), current L2 self (β = 0.285) and authenticity gap (β = 0.331). Accordingly, after applying and transferring AI feedback to target contexts, and evaluating its accuracy, learners' instrumentality-promotion to learn with AI and their current beliefs regarding what they currently learn, and achieved with it, increased. Additionally, they considered AILL might be more authentic than their previous language learning experiences. This was also addressed by S3 where she claimed that the information that AI provided was much more update for her, and other students which they did not know, and not covered in their previous language, and cultural courses. Several factors could explain these findings. The first one is that the both ideate and prototype were associated with learners' promotional instrumentality rather than the prevention one, in which they perceived AI feedback as more aligned with their personal goals than with extrinsic obligations, leading them to collect and transformed these feedback as required by the future personal objects, and target setting.

We also observed that students reported using AI to find the most accurate answers and the best sources of information it suggested. This may further explain the positive correlation between the test phase and learners' motivation, particularly the authenticity gap. This finding aligns

with previous studies indicating that design thinkers can transform their ideas into practical solutions through DE, especially during the test phase (Calavia et al., 2023; Yang et al., 2022). In summary, as learners progressed through the DEAILL phases, transitioning from divergent (Define) to convergent (Test) thinking, their motivation evolved accordingly—from prevention-focused to promotion-focused—and their current L2 self in AILL also developed.

5.3. Relationship between DEAILL and L2 grit

We also observed that students reported using AI to find the most accurate answers and the best sources of information it suggested. This may further explain the positive correlation between the test phase and learners' motivation, particularly the authenticity gap. This finding aligns with previous studies indicating that design thinkers can transform their ideas into practical solutions through DE, especially during the test phase (Calavia et al., 2023; Yang et al., 2022). In summary, as learners progressed through the DEAILL phases, transitioning from divergent (define) to convergent (test) thinking, their motivation evolved accordingly—from prevention-focused to promotion-focused—and their current L2 self in AILL also developed.

Moreover, Ideate skill positively shaped both language learners' consistency of efforts ($\beta = 0.215$) and persistency of interest ($\beta = 0.158$). Therefore, as language learners collected the information provided by AI based on their areas of learning needs and difficulties, their efforts and interest increased. This finding aligns with the study's qualitative results and is consistent with Brown's (2008) assertion that true design thinkers are driven by their ability to identify problems and explore potential solutions, which, in turn, fuels their determination to address those challenges.

Similarly, prototype skill predicted learners' L2 grit. Thus, the more language learners adapted AI feedback to their target context, the grittier they became. This may be attributed to the anthropomorphism of AI and its perceived intelligence, as well as its ability to generate novel ideas. Recent studies suggest that these characteristics encourage AI users to engage with the technology more frequently, because of AI's human-like behavior, which in turn shaped users' learning behaviors (Balakrishnan et al., 2022; Janson, 2023; Rahimi & Mosalli, 2025), including L2 grit. The qualitative findings of this study further confirm that learners perceived AI as both innovative and intelligent, which contributed to their sustained motivation.

The Test phase also significantly predicted both CI ($\beta = 0.247$) and PE ($\beta = 0.176$), suggesting that language learners' interest and efforts were

elevated and that they overcame any disappointments or challenges when learning language with AI. This resilience stemmed from their perception of AI-generated feedback and information as reliable once it was applied to the target context. These findings address gaps identified in previous studies, which have called for further exploration of the role of DT in shaping learners' psychological attributes (He et al., 2023; Tsai, 2021). Furthermore, they contribute to a deeper understanding of the factors influencing language learners' L2 grit. Recent studies in CALL have emphasized the importance of emotions (Chen Hsieh & Lee, 2021), motivation (Rahimi & Sevilla-Pavón, 2025a), autonomy (Paradowski & Jelińska, 2023), and mindset (Zarrinabadi et al., 2022) in this regard, and our findings build upon these insights by demonstrating the impact of the Test phase on Ci and PE. Additionally, this finding may serve as a response to concerns raised by Crompton and Burke (2024) regarding AI inaccuracies and hallucinations, which can lead to the spread of misinformation among learners. Through the Test phase, learners actively engaged in evaluating AI-generated feedback with their peers and teachers, assessing its reliability and relevance. This critical reflection process likely played a crucial role in fostering their L2 grit, resulting in the development of their L2 grit.

5.4. The mediation and moderation roles of AIL2MSS

An indirect bootstrap analysis of the PLS-SEM revealed some unexpected yet insightful findings regarding the serial correlations among DEAILL, AIL2MSS, and L2 grit. The serial mediation analysis indicated that only digital self-authenticity mediated the relationship between learners' Test skills (ID-AUT-CI) and their CI, as well as between Ideate skills and their CI (ID-AUT-CI).

Since both the direct (see Table 6) and indirect (see Table 7) effects were significant and aligned in the same direction, these mediations constitute a complementary type (Hair et al. 2023). As a result, we identified the sign of the authenticity gap in AILL and in Spain, a phenomenon previously reported in CALL research across various digital language learning environments, including LMOOCs (Rahimi 2023), virtual exchanges (Rahimi and Sevilla-Pavón 2025a), online language learning (Rahimi and Mosalli 2024; Smith et al. 2020), and digital language learning games (Henry and Cliffordson 2017) in the Iranian, Chinese, and Swedish contexts. This finding may be linked to the anthropomorphism of AI and its perceived intelligence. Prior research has shown that language learners (Balakrishnan et al. 2022; Janson 2023; Rahimi and Mosalli 2025) often attribute human-like qualities to AI tools, largely because of their ability to provide timely, objective feedback without emotional

biases and their capacity to deliver updated, contextually relevant information that aligns more closely with learners' academic and personal lives than traditional language instruction (Henry and Ushioda 2013Henry, 2013; Rahimi, 2023; Rahimi & Sevilla-Pavón, 2025a). Consequently, this complementary mediation suggests that the authenticity gap serves as a bridge between DEAILL and language learners' L2 grit, particularly influencing learners' consistency of interest, as illustrated in Figure 7.

Moreover, the remaining AIL2MSS factors did not function as mediator variables but rather as moderators, as illustrated in Figure 8. Specifically, while the Ideate, Prototype, and Test phases significantly predicted learners' L2 grit (see Table 8), the inclusion of instrumentality-promotion, instrumentality-prevention, and current L2-self altered these significant effects, rendering them insignificant (see Table 9). This indicates that these factors acted as moderator variables rather than mediators. In this regard, our findings contradict Feng and Papi's (2020) assertion that the ideal L2 self serves as a strong mediator in shaping learners' L2 grit. However, it is important to note that their

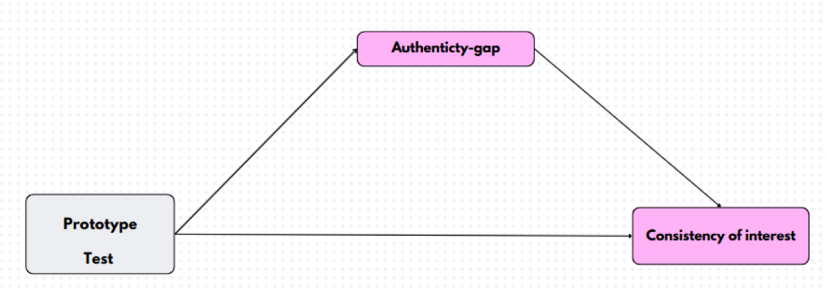

Figure 7. The mediation role of authenticy gap.

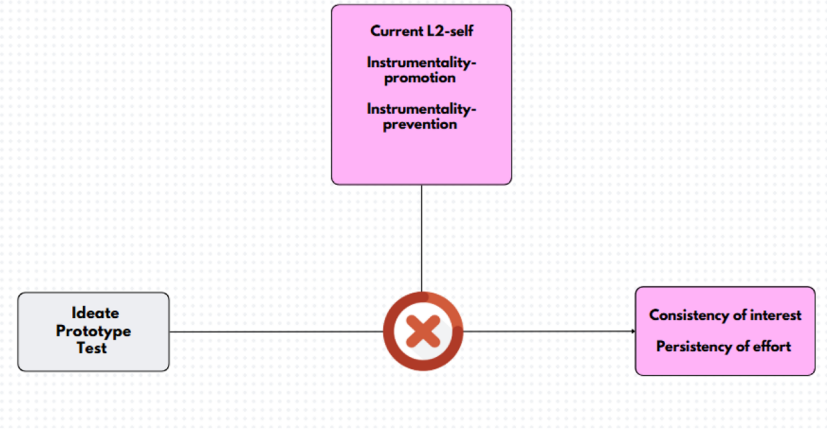

Figure 8. The moderation roles of the current L2 self, instrumentality-promotion, and instrumentality-prevention between ideation, Prototyping and Testing with learners' L2 grit.

study was conducted in the context of face-to-face language instruction. Our study expands upon these previous findings by emphasizing that digital self-authenticity serves as a stronger mediator of language learners' L2 grit in Artificial Intelligence Language Learning (AILL), while the other AIL2MSS factors function primarily as moderators. According to Al-Hoorie (2018), L2MSS is a strong predictor of learners' intended efforts, which depend on outcome variables. However, OL2MS did not emerge as an effective mediator in this study. Instead, our findings underscore the role of the authenticity gap as a significant mediator in the relationship between DEAILL and L2 grit.

Ultimately, the importance of motivation cannot be overstated, particularly in relation to L2 grit within digital language learning environments (Pawlak et al., 2022). Our findings highlight that digital self authenticity serves as a mediator between language learners' DEAILL skills and their L2 grit, emphasizing the transformative role of digital problem-solving skills in AI-assisted learning. This outcome aligns with the DEAILL framework, which is designed to develop creative problem solvers through the use of CALL materials, particularly AI.

Recent CALL research has stressed that face-to-face classrooms often offer limited opportunities for students to express themselves, think creatively, and develop autonomy. In contrast, authentic digital language learning contexts, particularly those outside formal schooling, have been shown to enhance learners' interest, attitude, and L2 grit (Rahimi, 2023), a pattern confirmed by our study.

Furthermore, Dörnyei and Ryan (2015) argue that language motivation is driven by the gap between a learner's current identity and their desired identity, encapsulated in the phrase "Don't dream it, be it." Expanding on this idea, we propose that in today's digital era, language motivation is shaped by the contrast between traditional language learning experiences (e.g., repetitive classroom activities) and the new skills acquired through AI and DEAILL (e.g., problem-solving and adaptive learning). This shift reflects a new perspective: "Don't repeat it, update it."

6. Conclusion and practical implications

The study The study validated the factorial structure of DEAILL within the CALL field and the AILL subfield. All DEAILL phases—except for the Define phase—significantly predicted language learners' current L2 self, instrumentality promotion, instrumentality prevention, and authenticity gap. Additionally, bootstrap direct analysis revealed that the Ideate, Prototype, and Test phases played a crucial role in shaping language learners' persistence of effort and consistency of interest, key L2 grit latent variables in AILL. Furthermore, our findings demonstrated that

digital self-authenticity mediated the relationship between DEAILL and consistency of interest in the AILL context, while current L2 self, instrumentality promotion, and instrumentality prevention acted as moderators in the relationship between DEAILL and L2 grit.

These findings contribute to CALL, AILL, Applied Linguistics, and Psycholinguistics in both theoretical and practical ways. First, the study introduces DEAILL as a new theoretical framework within CALL, developing, validating, and applying a novel instrument to assess language learners' design thinking skills in CALL, particularly in AILL. While previous research has validated design thinking in fields such as STEM (He et al., 2023), psychology (Tsai, 2021), and business (Chouyluam et al., 2021), this study specializes its application to CALL and AILL. Additionally, we are among the first to position DEAILL as a critical 21st-century digital skill shaping language learners' course progress in CALL, particularly from the perspective of complex dynamic systems (Larsen-Freeman & Cameron, 2008). While prior research has emphasized cognitive resources (Dörnyei & Ryan, 2015), personality traits (Teimouri et al., 2020), and psychological factors (Rahimi & Cheraghi, 2022), we now highlight 21st-century digital skills, particularly Design Thinking skills, as key factors enhancing motivation and L2 grit in AILL. Notably, while the Define phase only influenced instrumentality prevention, other phases—such as Ideation, Testing, and Prototyping—significantly strengthened instrumentality promotion, current L2 self, authenticity gap, consistency of interest, and persistence of effort in AILL.

the most necessary skills during this current age of innovation: design thinking skills and how they can shape learners' online motivation and L2 grit. Moreover, most of the studies have explored and applied L2 grit to predict performance in the classroom and psychological factors (e.g. Zhang, 2023), but we have shifted its role to be an outcome variable. Additionally, we totally shifted the view from exploring learners' psychological factors and their L2 grit to their digital problem-solving skills. This study is also noteworthy for exploring the authenticity gap in both AILL and Spain for the first time, and found that DEAILL is one of the predictors, a skill that cannot be acquired in traditional language learning.Thirdly, this study shifts the focus from traditional language skills or subskills taught with AI to the increasingly crucial design thinking skills and their role in shaping learners' online motivation and L2 grit in the current age of innovation. While most studies have explored L2 grit to predict classroom performance and psychological factors (e.g., Zhang, 2023), we have redefined its role as an outcome variable. This marks a significant departure from traditional approaches, as we also move away from focusing solely on learners' psychological

factors and L2 grit toward emphasizing their digital problem-solving skills. Furthermore, this study is noteworthy for exploring the authenticity gap in both AILL and Spain for the first time, identifying DEAILL as a key predictor—a skill that cannot be acquired through traditional language learning methods.

In alignment with previous research in language motivation and CALL, which called for the integration and validation of new L2MSS factors, namely current L2 self and the authenticity gap (Rahimi, 2023; Smith et al., 2020), we followed this recommendation. In addition, we have shifted the role of L2MSS from being a predictor variable (Rahimi, 2023) or outcome variable (Lamb & Arisandy, 2019) to functioning as both mediator and moderator variables in shaping L2 grit.

The findings of this study have significant practical implications at the macro, meso, and micro levels. Based on these results, several key recommendations are provided for language learners at the first level:

To cultivate a high level of L2 grit and foster success as design thinkers, language learners must have both professional and personal motivations for using artificial intelligence in their language learning. Before relying on AI, learners should first identify their weaknesses rather than simply accepting the information provided by AI tools. They should avoid directly applying AI-generated content; instead, they should engage in a process of gathering ideas, brainstorming, exchanging feedback, cross-referencing sources, and selecting the most relevant and appropriate information before applying it to their learning environment. We recommend that learners assess the direct and indirect responses from others in the target context to evaluate the reliability of AI feedback. Our quantitative and qualitative findings revealed that during the initial phase of DT, Define, learners with a high degree of divergence tended to focus on instrumentality-prevention. This indicates that they primarily sought to address academic problems with AI, rather than pursuing broader L2 self-guides, such as instrumentality-promotion or their current L2 self. To support the development of learners' full range of L2 self-identities, language teachers should encourage students to think beyond simply passing a course while using AI. Instead, they should be prompted to consider their future goals and other current learning needs that AI might help alleviate before utilizing it. Moreover, learners found that the ideas and feedback provided by AI were more authentic, updated, and rational compared to those encountered in previous language learning contexts. These insights aligned with their future ideal selves and current needs, and significantly acted as a mediator between the learners' DEAILL skills and their L2 grit.

The second level focuses on language teachers. Rather than solely using technology or concentrating on students' final scores, instructors

should prioritize addressing language learners' challenges, particularly through the use of AI. To align with the demands of the new generation, learners must become successful problem solvers as well as proficient language users in this era of rapid innovation (Rahimi, 2024b; Rahimi & Sevilla-Pavón, 2024). Therefore, language teachers must foster both language proficiency and problem-solving skills simultaneously. In today's world, developing design thinkers is just as crucial as developing proficient language speakers. Language teachers should not simply integrate AI tools and leave students to navigate them on their own; instead, they should guide students in learning how to use AI effectively for language learning, particularly within the framework of design thinking. Teachers should take into account students' personal and professional goals (IPO), external obligations (IPR), and areas of difficulty, and then use AI strategically to address these needs. By guiding learners through the DEAILL steps and principles, teachers can help them not only enhance their design thinking skills but also increase their motivation and effort to learn the language with AI.

Pedagogical experts, CALL administrators, and AI designers should encourage teachers to incorporate problem-solving strategies, particularly within the context of CALL and AI, with a focus on addressing language learners' challenges related to their future and current L2 selves and instrumentality-prevention. This approach will positively influence learners' L2 grit and motivation for language learning. One effective way to achieve this is by allocating additional funds for workshops that train language teachers on how to integrate DEAILL into their teaching practices. AI designers should also develop AI-based activities specifically tailored to the CALL field, taking into account learners' L2 identities and the DEAILL framework. Furthermore, CALL specialists and CEFR experts, along with executive managers in the Ministry of Education, should familiarize themselves with DEAILL as a language theory and incorporate it as a new criterion and metric when evaluating language teachers. This integration will ensure that DEAILL becomes a fundamental aspect of language teaching and learning in the digital age.

With the development of the DEAILL theory and its associated instructions, we have provided stakeholders, teachers, and scholars with a new opportunity to move beyond traditional language skills and embrace 21st-century digital competences. This framework opens up avenues for testing, extending, validating, and replicating in various language contexts and CALL fields. The theory and methods presented can also serve as a response to concerns raised in previous studies, which suggested that artificial intelligence might reduce learners' higher-order thinking, hinder contextual understanding (Farrokhnia et

al., 2023), perpetuate discrimination (Farrokhnia et al., 2023), and exacerbate biases in education. Despite these challenges, it is essential for educators, researchers, and teachers to collaboratively explore strategies for mitigating these drawbacks, rather than merely emphasizing them. We hope that DEAILL offers a practical approach to address these concerns.With the development of the DEAILL theory and its associated instructions, we have provided stakeholders, teachers, and scholars with a new opportunity to move beyond traditional language skills and embrace 21st-century digital competences. This framework opens up avenues for testing, extending, validating, and replicating in various language contexts and CALL fields. The theory and methods presented can also serve as a response to concerns raised in previous studies, which suggested that artificial intelligence might reduce learners' higher-order thinking, hinder contextual understanding (Farrokhnia et al., 2023), perpetuate discrimination (Farrokhnia et al., 2023), and exacerbate biases in education. Despite these challenges, it is essential for educators, researchers, and teachers to collaboratively explore strategies for mitigating these drawbacks, rather than merely emphasizing them. We hope that DEAILL offers a practical approach to address these concerns.

7. Future directions

Our research successfully validates the factorial structure of DEAILL in CALL, AILL, and Spanish EFL contexts, and demonstrates its role in shaping AIL2MSS and L2 grit. However, to further advance the CALL literature, several limitations must be considered. First, the DEAILL framework and its questionnaire need to be replicated and validated across other EFL and ESL contexts. Additionally, the sample size in this study was limited, and there was a gender imbalance, which should be addressed in future research. Further qualitative studies, such as observations or mapping exercises, could provide deeper insights into how DEAILL influences learners' AIL2MSS and grit. Future research could also employ longitudinal approaches to investigate the authenticity gap between AILL and previous language learning experiences through the lens of Design Thinking skills in AILL. Additionally, CALL researchers are encouraged to explore the pedagogical, technological, and psychological factors that shape learners' DEAILL skills. Lastly, extending this research to examine the impact of DEAILL on other psychological and pedagogical factors in CALL could provide valuable contributions to the field.To begin, our research validates the factorial structure of DEAILL in CALL, AILL, and Spanish EFL and its role in shaping AIL2MSS and L2 grit. In order to move forward with

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Acknowledgements are due to the Spanish Ministry of Science and Innovation for funding the research project SOCIOEMOVE: DEVELOPING SOCIOEMOTIONAL SKILLS THROUGH VIRTUAL EXCHANGE (SOCIOEMOVE: DESARROLLO DE HABILIDADES SOCIOEMOCIONALES EN INTERCAMBIOS VIRTUALES), Ref. PID2023-151087NB-I00, "Knowledge Generation Projects" call from the State Programme to Promote Scientific-Technical Research and its Transfer, within the framework of the State Plan for Scientific, Technical Research and Innovation 2021-2023, co-directed by Dr Ana Sevilla-Pavón and Dr Margarita Vinagre-Laranjeira from 01/09/2024 to 01/09/2027.

ORCID

Ana Sevilla-Pavón http://orcid.org/0000-0001-5253-0423

References

Alamer, A. (2021). Grit and language learning: Construct validation of L2-Grit scale and its relation to later vocabulary knowledge. *Educational Psychology*, *41*(5), 544–562. https://doi.org/10.1080/01443410.2020.1867076

Al-Hoorie, A. H. (2018). The L2 motivational self-system: A meta-analysis. *Studies in Second Language Learning and Teaching*, *8*(4), 721–754. https://doi.org/10.14746/ssllt.2018.8.4.2

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*, 1–22. https://doi.org/10.1080/09588221.2023.2246519

Balakrishnan, B. (2021). Exploring the impact of design thinking tool among design undergraduates: A study on creative skills and motivation to think creatively. *International Journal of Technology and Design Education*, *32*(3), 1799–1812. https://doi.org/10.1007/s10798-021-09652-y

Balakrishnan, J., Abed, S. S., & Jones, P. (2022). The role of meta-UTAUT factors, perceived anthropomorphism, perceived intelligence, and social self-efficacy in chatbot-based services? *Technological Forecasting and Social Change*, *180*, 121692. https://doi.org/10.1016/j.techfore.2022.121692

Brown, T. (2008). Design thinking. *Harvard Business Review*, *86*, 84–92.

Brown, T., & Katz, B. (2011). Change by design. *Journal of Product Innovation Management*, *28*(3), 381–383. https://doi.org/10.1111/j.1540-5885.2011.00806.x

Brown, T., & Wyatt, J. (2010). Design thinking for social innovation. *Development Outreach*, *12*(1), 29–43. https://doi.org/10.1596/1020-797X_12_1_29

Calavia, M. B., Blanco, T., Casas, R., & Dieste, B. (2023). Making design thinking for education sustainable: Training preservice teachers to address practice challenges. *Thinking Skills and Creativity*, *47*, 101199. https://doi.org/10.1016/j.tsc.2022.101199

Changlek, A., & Palanukulwong, T. (2015). Motivation and Grit: Predictors of language learning achievement. *Veridian E-Journal, Silpakorn University (Humanities, Social Sciences and Arts)*, *8*(4), 23–36.

Chen Hsieh, J., & Lee, J. S. (2021). Digital storytelling outcomes, emotions, grit, and perceptions among EFL middle school learners: Robot-assisted versus PowerPoint-assisted presentations. *Computer Assisted Language Learning*, *36*(5–6), 1088–1115. https://doi.org/10.1080/09588221.2021.1969410

Chouyluam, S., Wannapiroon, P., & Nilsook, P. (2021). Creative design thinking learning model integrated immersive experiential marketing to enhance digital entrepreneurs. *International Journal of Trade, Economics and Finance*, *12*(1), 26–32. https://doi.org/10.18178/ijtef.2021.12.1.689

Cleminson, T., & Cowie, N. (2021). Using design thinking as an approach to creative and communicative engagement in the English as a Foreign Language (EFL) classroom. *Journal of University Teaching and Learning Practice, 18*(4), 7. https://doi.org/10.53761/1.18.4.7

Colpaert, J. (2020). Editorial position paper: How virtual is your research? *Computer Assisted Language Learning*, *33*(7), 653–664. https://doi.org/10.1080/09588221.2020.1824059

Credé, M., Tynan, M. C., & Harms, P. D. (2017). Much ado about grit: A meta-analytic synthesis of the grit literature. *Journal of Personality and Social Psychology, 113*(3), 492–511. https://doi.org/10.1037/pspp0000102

Creswell, J. W., & Clark, V. L. P. (2011). *Designing and conducting mixed methods research*. SAGE.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. SAGE.

Crites, K., & Rye, E. (2020). Innovating language curriculum design through design thinking: A case study of a blended learning course at a Colombian university. *System, 94*, 102334. https://doi.org/10.1016/j.system.2020.102334

Crompton, H., & Burke, D. (2024). The educational affordances and challenges of ChatGTP: State of the field. *TechTrends*, *68*(2), 380–392. https://doi.org/10.1007/s11528-024-00939-0

Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. Routledge.

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology, 92*(6), 1087–1101. https://doi.org/10.1037/0022-3514.92.6.1087

Duckworth, A. L., & Quinn, P. D. (2009). Development and validation of the short grit scale (grit–s). *Journal of Personality Assessment*, *91*(2), 166–174. https://doi.org/10.1080/00223890802634290

Dym, C. L., Agogino, A. M., Eris, O., Frey, D. D., & Leifer, L. J. (2005). Engineering design thinking, teaching, and learning. *Journal of Engineering Education*, *94*(1), 103–120. https://doi.org/10.1002/j.2168-9830.2005.tb00832.x

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, *61*(3), 460–474. https://doi.org/10.1080/14703297.2023.2195846

Feng, L., & Papi, M. (2020). Persistence in language learning: The role of grit and future self-guides. *Learning and Individual Differences, 81*, 101904. https://doi.org/10.1016/j.lindif.2020.101904

Gimeno-Sanz, A. (2015). Moving a step further from "integrative CALL". What's to come? *Computer Assisted Language Learning*, *29*(6), 1102–1115. https://doi.org/10.1080/09588221.2015.1103271

Hair, J. F., Jr., Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2023). *Advanced issues in partial least squares structural equation modeling*. SAGE Publications.

Hair, J. F., Risher, J. J., Sarstedt, M., & Ringle, C. M. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, *31*(1), 2–24. https://doi.org/10.1108/EBR-11-2018-0203

Hays, L., Jurkowski, O., & Sims, S. K. (2024). ChatGPT in K-12 education. *TechTrends*, *68*(4), 281–294. https://doi.org/10.1007/s11528-023-00924-z

He, W., Yan, J., Wang, C., Liao, L., & Hu, X. (2023). Exploring the impact of the design thinking model on fifth graders' creative self-efficacy, situational interest, and individual interest in STEM education. *Thinking Skills and Creativity*, *50*, 101424. https://doi.org/10.1016/j.tsc.2023.101424

Henry, A. (2013). Digital games and ELT: Bridging the authenticity gap. In E. Ushioda (Ed.), *International perspectives on motivation: Language learning and professional challenges* (pp. 133–155). Springer. https://doi.org/10.1057/9781137000873

Henry, A., & Cliffordson, C. (2017). The impact of out-of-school factors on motivation to learn English: Self-discrepancies, beliefs, and experiences of self-authenticity. *Applied Linguistics*, *38*(5), 713–736. https://doi.org/10.1093/applin/amv060

Henseler, J., Ringle, C. M., & Sarstedt, M. (2014). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, *43*(1), 115–135. https://doi.org/10.1007/s11747-014-0403-8

Hews, R., Beligatamulla, G., & McNamara, J. (2023). Creative confidence and thinking skills for lawyers: Making sense of design thinking pedagogy in legal education. *Thinking Skills and Creativity*, *49*, 101352. https://doi.org/10.1016/j.tsc.2023.101352

Janson, A. (2023). How to leverage anthropomorphism for chatbot service interfaces: The interplay of communication style and personification. *Computers in Human Behavior*, *149*, 107954. https://doi.org/10.1016/j.chb.2023.107954

Jeon, J. (2022). Exploring AI chatbot affordances in the EFL classroom: Young learners' experiences and perspectives. *Computer Assisted Language Learning*, *37*(1–2), 1–26.https://doi.org/10.1080/09588221.2021.2021241

Kimbell, L. (2011). Rethinking design thinking: Part I. *Design and Culture*, *3*(3), 285–306. https://doi.org/10.2752/175470811X13071166525216

Lake, J, D. Da Silva (2013). Positive L2 self: Linking positive psychology with L2 motivation. In M. T. Apple, & T. Fellner (Eds.), *Language learning motivation in Japan.* (pp. 225–244) Multilingual Matters.

Lam, K. K. L., & Zhou, M. (2022). Grit and academic achievement: A comparative cross-cultural meta-analysis. *Journal of Educational Psychology*, *114*(3), 597–621. https://doi.org/10.1037/edu0000699

Lamb, M., & Arisandy, F. E. (2019). The impact of online use of English on motivation to learn. *Computer Assisted Language Learning*, *33*(1–2), 85–108. https://doi.org/10.1080/09588221.2018.1545670

Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. OUP Oxford.

Leem, S., & Lee, S. W. (2024). Fostering collaboration and interactions: Unveiling the design thinking process in interdisciplinary education. *Thinking Skills and Creativity*, *52*, 101520. https://doi.org/10.1016/j.tsc.2024.101520

Li, C., & Yang, Y. (2023). Domain-general grit and domain-specific grit: Conceptual structures, measurement, and associations with the achievement of German as a foreign language. *International Review of Applied Linguistics in Language Teaching*, *62*(4), 1513–1537. https://doi.org/10.1515/iral-2022-0196

Lin, M.-Y., & Chang, Y.-S. (2024). Using design thinking hands-on learning to improve artificial intelligence application creativity: A study of brainwaves. *Thinking Skills and Creativity*, *54*, 101655. https://doi.org/10.1016/j.tsc.2024.101655

Liu, S., & Li, C. (2023). Promoting design thinking and creativity by making: A quasi-experiment in the information technology course. *Thinking Skills and Creativity*, *49*, 101335. https://doi.org/10.1016/j.tsc.2023.101335

Melendy, G. (2008). Motivating writers: The power of choice. *The Asian EFL Journal*, *10*(3), 187–198.

Mercer, S. (2018). Psychology for language learning: Spare a thought for the teacher. *Language Teaching*, *51*(4), 504–525. https://doi.org/10.1017/S0261444817000258

Micheli, P. G., Bhatti, S., Wilner, S., & Beverland, M. (2018). Doing design thinking: Conceptual review, synthesis and research agenda. *Academy of Management Proceedings*, *2018*(1), 16071. https://doi.org/10.5465/AMBPP.2018.16071abstract

Nazim, M., & Mohammad, T. (2022). Implications of design thinking in an EFL classroom: Writing in context. *Theory and Practice in Language Studies, 12*(12), 2723–2730. https://doi.org/10.17507/tpls.1212.31

Opincāne, M., Laganovska, K. (2023). Applying of design thinking approach in the development of foreign language course. *Society Integration Education. Proceedings of the International Scientific Conference, 1*, 155–164. https://doi.org/10.17770/sie2023vol1.7102

Outeda, C. C. (2024). European education area and digital education action plan (2021–2027): One more step towards the europeanisation of education policy. In D. Ramiro Troitiño (Ed.), *E-Governance in the European Union. Contributions to political science*. Springer. https://doi.org/10.1007/978-3-031-56045-3_13

Paradowski, M. B., & Jelińska, M. (2023). The predictors of L2 grit and their complex interactions in online foreign language learning: Motivation, self-directed learning, autonomy, curiosity, and language mindsets. *Computer Assisted Language Learning*, 1–38. https://doi.org/10.1080/09588221.2023.2192762

Pawlak, M., Chengchen, L., Zawodniak, J., & Kruk, M. (2024). Un análisis de los efectos predictivos de la tenacidad general y la tenacidad sobre la L2 en la motivación: El efecto mediador del conocimiento auto-percibido de la L2. *Porta Linguarum* , *XXIX*, 1–18. https://doi.org/10.30827/portalin.viIX.29884

Pawlak, M., Csizér, K., Kruk, M., & Zawodniak, J. (2022). Investigating grit in second language learning: The role of individual difference factors and background variables. *Language Teaching Research*, https://doi.org/10.1177/13621688221105775

Rahimi, A. R. (2023). The role of EFL learners' L2 self-identities, and authenticity gap on their intention to continue LMOOCs: Insights from an exploratory partial least approach. *Computer Assisted Language Learning*, 1–32. https://doi.org/10.1080/09588221.2023.2202215

Rahimi, A. R. (2024a). A bi-phenomenon analysis to escalate higher educators' competence in developing university students' information literacy (HECDUSIL): The role of language lectures' conceptual and action-oriented digital competencies and skills. *Education and Information Technologies*, *29*(6), 7195–7222. https://doi.org/10.1007/s10639-023-12081-0

Rahimi, A. R. (2024b). Beyond digital competence and language teaching skills: The bi-level factors associated with EFL teachers' 21st-century digital competence to cultivate 21st-century digital skills. *Education and Information Technologies*, *29*, 9061–9089. https://doi.org/10.1007/s10639-023-12171-z

Rahimi, A. R., & Cheraghi, Z. (2022). Unifying EFL learners' online self-regulation and online motivational self-system in MOOCs: A structural equation modeling approach. *Journal of Computers in Education*, *11*, 1–27. https://doi.org/10.1007/s40692-022-00245-9

Rahimi, A. R., & Mosalli, Z. (2024). Exploring the direct and indirect effects of EFL learners' online motivational self-system on their online language learning acceptance: The new roles of current L2 self and digital self-authenticity. *Asian-Pacific Journal of*

Second and Foreign Language Education, *9*(1), 49. https://doi.org/10.1186/s40862-024-00266-0

Rahimi, A. R., & Mosalli, Z. (2025). Language learners' surface, deep, and organizing approaches to ChatGPT-assisted language learning: What contextual, individual, and ChatGPT-related factors contribute? *Smart Learning Environments*, *12*, 16. https://doi.org/10.1186/s40561-025-00368-3

Rahimi, A. R., & Sevilla-Pavón, A. (2024). The role of ChatGPT readiness in shaping language teachers' language teaching innovation and meeting accountability: A bisymmetric approach. *Computers and Education: Artificial Intelligence*, *7*, 100258. https://doi.org/10.1016/j.caeai.2024.100258

Rahimi, A. R., & Sevilla-Pavón, A. (2025a). Modeling the relationship between online L2 motivational selfsystem and EFL learners' virtual exchange self-regulations: The mediator and moderator roles of L2 grit. *ReCALL*, 1–20. https://doi.org/10.1017/S095834402400034X

Rahimi, A. R., & Sevilla-Pavón, A. (2025b). Scaling up computational thinking skills in computer-assisted language learning (CTsCALL) and its fitness with language learners' intentions to use virtual exchange: A bi-symmetric approach. *Computers in Human Behavior Reports*, *17*, 100607. https://doi.org/10.1016/j.chbr.2025.100607

Rahimi, A. R., Sheyhkholeslami, M., & Mahmoudi Pour, A. (2025). Uncovering personalized L2 motivation and self-regulation in ChatGPT-assisted language learning: A hybrid PLS-SEM-ANN approach. *Computers in Human Behavior Reports*, *17*, 100539. https://doi.org/10.1016/j.chbr.2024.100539

Roberts, Y. (2009). *Grit: The skills for success and how they are grown*. Young foundation.

Rowe, P. G. (1991). *Design thinking*. MIT Press.

Sarstedt, M., Hair, J. F., Jr, Nitzl, C., Ringle, C. M., & Howard, M. C. (2020). Beyond a tandem analysis of SEM and PROCESS: Use of PLS-SEM for mediation analyses!. *International Journal of Market Research*, *62*(3), 288–299. https://doi.org/10.1177/1470785320915686

Sevilla-Pavón, A., & Nicolaou, A. (2020). Artefact co-construction in virtual exchange: 'Youth Entrepreneurship for Society. *Computer Assisted Language Learning*, *35*(7), 1642–1667. https://doi.org/10.1080/09588221.2020.1825096

Simon, H. A. (2019). *The sciences of the artificial*. The MIT Press. https://doi.org/10.7551/mitpress/12107.001.0001

Smith, S. A., Foster, M. E., Baffoe-Djan, J. B., Li, Z., & Yu, S. (2020). Unifying the current self, ideal self, attributions, self-authenticity, and intended effort: A partial replication study among Chinese university English learners. *System*, *95*, 102377. https://doi.org/10.1016/j.system.2020.102377

Sotlikova, R. (2023). Design thinking in education: Empowering students in ELT class. *Proceedings Series on Social Sciences & Humanities, 13*, 196–199. https://doi.org/10.30595/pssh.v13i.904

Sreenivasan, A., & Suresh, M. (2024). Design thinking and artificial intelligence: A systematic literature review exploring synergies. *International Journal of Innovation Studies*, *8*(3), 297–312. https://doi.org/10.1016/j.ijis.2024.05.001

Stanford D. School (2010). *An introduction to design thinking PROCESS GUIDE*. Institute of Design at Stanford.

Staub, L., van Giffen, B., Hehn, J., et al. (2023). Design thinking for artificial intelligence: How design thinking can help organizations to address common AI project challenges. In *Lecture Notes in Computer Science* (pp. 251–267). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-48057-7_16

Teimouri, Y., Plonsky, L., & Tabandeh, F. (2022). L2 grit: Passion and perseverance for second-language learning. *Language Teaching Research, 26*(5), 893–918. https://doi.org/10.1177/1362168820921895

Tsai, C.-A., Song, M.-Y. W., Lo, Y.-F., & Lo, C.-C. (2023). Design thinking with constructivist learning increases the learning motivation and wicked problem-solving capability—An empirical research in Taiwan. *Thinking Skills and Creativity*, *50*, 101385. https://doi.org/10.1016/j.tsc.2023.101385

Tsai, M.-J., & Wang, C.-Y. (2020). Assessing young students' design thinking disposition and its relationship with computer programming self-efficacy. *Journal of Educational Computing Research*, *59*(3), 410–428. https://doi.org/10.1177/0735633120967326

Wang, X., Pang, H., Wallace, M. P., Wang, Q., & Chen, W. (2022). Learners' perceived AI presences in AI-supported language learning: A study of AI as a humanized agent from community of inquiry. *Computer Assisted Language Learning*, 1–27. https://doi.org/10.1080/09588221.2022.2056203

Wei, H., Gao, K., & Wang, W. (2019). Understanding the relationship between grit and foreign language performance among middle school students: The roles of foreign language enjoyment and classroom environment. *Frontiers in Psychology*, *10*,1508. https://doi.org/10.3389/fpsyg.2019.01508

Yamashita, T. (2018). *Grit and second language acquisition: Can passion and perseverance predict performance in Japanese language learning?* [Unpublished MA thesis]. University of Massachusetts.

Yang, X., Zhang, M., Zhao, Y., Wang, Q., & Hong, J.-C. (2022). Relationship between creative thinking and experimental design thinking in science education: Independent or related. *Thinking Skills and Creativity*, *46*, 101183. https://doi.org/10.1016/j.tsc.2022.101183

You, C., & Dörnyei, Z. (2015). Language learning motivation in China: Results of a large-scale stratified survey. *Applied Linguistics*, *37*(4), 495–519. https://doi.org/10.1093/applin/amu046

Yousefi, M., & Mahmoodi, M. H. (2022). The L2 motivational self-system: A meta-analysis approach. *International Journal of Applied Linguistics*, *32*(2), 274–294. https://doi.org/10.1111/ijal.12416

Zarrinabadi, N., Rezazadeh, M., & Mohammadzadeh Mohammadabadi, A. (2022). L2 grit and language mindsets as predictors of EFL learners' attitudes toward effectiveness and value of CALL. *Computer Assisted Language Learning*, *37*(7), 1726–1743. https://doi.org/10.1080/09588221.2022.2108061

Zhang, J., & Zhang, L. J. (2023). Examining the relationship between English as a foreign language learners' cognitive abilities and L2 grit in predicting their writing performance. *Learning and Instruction*, *88*, 101808. https://doi.org/10.1016/j.learninstruc.2023.101808

Zhao, X., & Wang, D. (2023). Domain-specific L2 grit, anxiety, boredom, and enjoyment in online chinese learning. *The Asia-Pacific Education Researcher*, 33, 783–794. https://doi.org/10.1007/s40299-023-00777-3

Zheng, C., Liang, J.-C., Li, M., & Tsai, C.-C. (2018). The relationship between English language learners' motivation and online self-regulation: A structural equation modelling approach. *System, 76*, 144–157. https://doi.org/10.1016/j.system.2018.05.003

Appendix A

Table A1. The study scale with descriptive statistics

Latent variable	The questions	Mean	Standard deviation	Excess kurtosis	Skewness	Cramér-von Mises test statistic	Cramér-von Mises p value
IPR1	Learning English with artificial intelligence is essential for me as this language can make a difference in my life in the future (Adapted from Rahimi, 2023)	3.304	0.869	−0.436	−0.135	0.922	0.000
IPR2	Learning English with Artificial Intelligence is essential for me because acquiring the language will change my life in the future	3.402	0.921	−0.834	0.039	0.788	0.000
IPR3	Learning English with artificial intelligence is important to me because an educated person is supposed to be able to speak English	3.391	0.807	−0.534	−0.081	1.074	0.000
IPO1	I have to learn English with artificial intelligence so as to pass this course (Adapted from Rahimi, 2023)	3.413	1.085	−0.934	−0.058	0.581	0.000
IPO2	I have to learn English with artificial intelligence to pass my final semester	3.402	1.094	−0.629	−0.151	0.606	0.000
IPO3	When thinking of not becoming a successful language learner with artificial intelligence, I feel scared	3.457	1.088	−0.532	−0.119	0.805	0.000
AUT1	Learning English with artificial intelligence is more interesting than a face-to-face language class (Adapted from Rahimi, 2023)	3.337	0.838	0.553	−0.597	1.146	0.000
AUT2	Learning English with artificial intelligence motivates me more than a face-to-face course	3.380	0.895	−0.305	0.090	0.959	0.000
AUT3	The actual process of English language learning with artificial intelligence feels more real compared to the things we do in traditional language classes	3.391	0.920	−0.220	−0.269	0.857	0.000
CU1	Now, I see myself as someone who is good at learning English with artificial intelligence (Adapted from Smith et al., 2020)	3.359	0.928	−0.398	−0.366	0.940	0.000
CU2	Being someone who can learn English with artificial intelligence is part of the person I am now	3.304	0.964	−0.841	−0.057	0.928	0.000
CU3	Now, I see myself as someone who can use artificial intelligence to learn language in many different situations	3.424	0.875	−0.347	−0.356	1.046	0.000
DE1	I usually see the gap between my current language needs and problems, and the goal which might be achieved through artificial intelligence (Adapted from Tsai & Wang, 2020).	3.283	0.838	−0.184	−0.017	1.027	0.000
DE2	I usually try to clarify my language learning problems that must be solved through artificial intelligence	3.163	1.076	−0.595	−0.120	0.589	0.000
DE3	I usually try to clearly identify my current and future language needs that can be met by using artificial intelligence	3.120	0.931	−0.543	0.086	0.779	0.000
ID1	I consciously seek out helpful suggestions and information that are generated by artificial intelligence to solve my language problems (Adapted from Tsai & Wang, 2020).	2.641	0.984	−0.588	0.153	0.713	0.000
ID2	I brainstorm the constructive ideas and information that are generated by artificial intelligence in my language learning.	2.761	0.971	−0.764	0.139	0.792	0.000

(*Continued*)

Table A1. Continued.

Latent variable	The questions	Mean	Standard deviation	Excess kurtosis	Skewness	Cramér-von Mises test statistic	Cramér-von Mises p value
ID3	I carefully collect accurate feedback and information generated by artificial intelligence in my language learning.	2.728	1.012	−0.671	0.188	0.718	0.000
PR1	I usually apply artificial intelligence feedback that I collect to my language tasks (Adapted from Tsai & Wang, 2020).	3.304	1.149	−0.844	−0.317	0.699	0.000
PR2	I usually develop an idea/solution based on artificial intelligence feedback	3.196	0.958	−0.546	−0.179	0.770	0.000
PR3	In order to meet my language needs, I usually utilize artificial intelligence feedback that I believe to be accurate and relevant	3.239	0.993	−0.693	−0.230	0.794	0.000
TE1	I evaluate the information provided by artificial intelligence based on the feedback provided by my teacher in my language learning class (Adapted from Tsai & Wang, 2020).	3.413	0.836	−0.053	−0.117	1.040	0.000
TE2	I evaluate the validity of the information provided by artificial intelligence based on the accuracy with which it is used in my target language setting and the environment where it is used	3.500	0.866	−0.636	0.153	0.968	0.000
TE3	I evaluate the validity of information provided by artificial intelligence based on the accuracy and its fitness with which it is used in my target language setting and the environment where it is used	3.457	0.914	−0.437	−0.173	0.840	0.000
CI1	My interest in learning English has changed from the time that I first used Artificial Intelligence for language learning (Adapted and modified from Teimouri et al., 2020).	3.478	0.814	0.131	−0.236	1.110	0.000
CI2	Since we used artificial intelligence for language learning, I have lost my interest in learning a language in a traditional setting	3.489	0.891	−0.284	−0.107	0.892	0.000
CI3	I am not as interested in learning a foreign language in traditional settings as I used to be	3.500	0.878	0.249	−0.391	0.985	0.000
PE1	Now that I have decided to learn English with artificial intelligence, nothing can prevent me from reaching this goal (Adapted and modified from Teimouri et al., 2020).	3.413	0.874	−0.271	−0.126	0.925	0.000
PE2	When it comes to learning language with artificial intelligence, I am a hard-working learner	3.315	0.932	−0.561	0.063	0.792	0.000
PE3	I will not allow anything to stop me from my progress in learning a language with artificial intelligence.	3.239	0.993	−0.483	−0.297	0.765	0.000