Current Psychology (2026) 45:1810
https://doi.org/10.1007/s12144-026-10370-8

Digital self-expression, self-efficacy, and foreign language enjoyment in an avatar-mediated metaverse listening environment

Esen Genc[1]

Received: 19 June 2026 / Accepted: 12 November 2026
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2026

Abstract
Foreign language enjoyment is associated with engagement and achievement in language learning yet little is known about how it relates to self-efficacy and motivation in immersive, digital environments. Drawing on Social Cognitive Theory, Self-Determination Theory, and Control-Value Theory, this study examined relationships among self-efficacy, motivation, and enjoyment in a metaverse listening environment, alongside learners' perceptions of avatar-based digital self-expression. Using a convergent mixed-methods design, 131 intermediate-level Turkish EFL undergraduates created personalized avatars and completed a 10-minute avatar-mediated listening task in a virtual café on Spatial.io. Participants then completed measures of self-efficacy, motivation, and enjoyment and answered open-ended questions. Because all variables were measured simultaneously, indirect effects are interpreted as associations consistent with the hypothesized model. Bias-corrected bootstrapping showed that self-efficacy was associated with enjoyment directly ($\beta = 0.47$) and indirectly through motivation ($\beta = 0.25$, 95% CI [0.14, 0.39]), explaining 66% of enjoyment variance. Sensitivity analyses indicated that alternative directional models fit equally well. Fuzzy-set qualitative comparative analysis further showed that high self-efficacy combined with high motivation was sufficient for high enjoyment under every calibration examined, whereas their joint absence was sufficient for low enjoyment only under distribution-based calibrations. Qualitative analysis identified eleven codes, with avatar-based digital self-expression and a reported sense of competence most frequent. Learners' accounts linked digital self-expression to both direct and motivation-related pathways, offering directions for experimental research and the design of avatar-mediated language-learning environments.

Keywords Metaverse · Digital self-expression · Self-efficacy · Foreign language enjoyment · Avatar · English as a foreign language

Introduction

Motivation, enjoyment, and self-efficacy are closely linked in the study of English as a foreign language (EFL), and immersive technologies such as the metaverse are changing how learners experience all these three. This shift is part of a wider affective turn in language education, which treats emotions and beliefs, including enjoyment and self-efficacy, as central to learning outcomes (Dewaele et al., 2019; Wu & Kabilan, 2025). Yet little is known about how self-efficacy, motivation, and enjoyment interact inside immersive, avatar-mediated environments specifically, or about how the technology itself shapes that interaction. Most existing work on these three constructs, moreover, treats them separately, using different theories and different methods.

Motivation in foreign language learning refers to the internal drive and goal-directed effort that learners invest in acquiring a second or foreign language, encompassing both the desire to engage with the learning process and the persistence to sustain that engagement over time (Deci & Ryan, 2000; Dewaele & Meftah, 2024). It draws on both intrinsic sources (e.g., personal interest, the ideal L2 self) and extrinsic sources (e.g., career aspirations, the ought-to self). Prior research has consistently shown that a robust ideal L2 self is associated with motivation to engage in both formal and informal digital learning contexts. Not all motivation, however, fosters enjoyment or enhances self-efficacy: promotion-focused motivation (the desire to advance) tends to predict positive outcomes more reliably

✉ Esen Genc
esen.genc@omu.edu.tr

[1] Ondokuz Mayıs University, Samsun, Türkiye

Published online: 20 November 2026

than prevention-focused motivation (the desire to avoid failure) (An et al., 2021; Liu et al., 2024).

Foreign language enjoyment (FLE) is defined as a complex, multidimensional positive emotion that learners experience while learning and using a foreign language, encompassing feelings of pleasure, intellectual stimulation, and social connection that arise in the language classroom and beyond (Dewaele & MacIntyre, 2014; Sheng et al., 2025). As the most clearly identifiable positive emotion in L2 learning, it is widely regarded as an important predictor of sustained engagement and learning outcomes. Empirical research has consistently shown that FLE is positively associated with learner engagement, motivation, and academic achievement in EFL contexts (Li & Wang, 2024). Enjoyment not only promotes direct participation in learning activities but also operates as a mediating mechanism through which motivation and self-efficacy translate into actual learning behavior (An et al., 2021; Liu et al., 2024). Enjoyment and motivation further reinforce each other, particularly as learners become more proficient: enjoyment elicits motivation to learn, and motivation in turn elicits enjoyment, creating a virtuous cycle of engagement and achievement (Dewaele & Meftah, 2024; Saito et al., 2025).

Self-efficacy, rooted in Bandura's (1997) Social Cognitive Theory, refers to an individual's belief in their capacity to execute the behaviors required to produce a given outcome, and in the EFL context, it captures learners' confidence in their ability to comprehend, produce, and engage with the target language across a range of tasks (Bandura, 1997; Li & Wang, 2024; Y. Wang, Shen et al., 2021). As the belief that one can successfully accomplish a task, self-efficacy both drives and results from motivation and enjoyment. Learners with high self-efficacy tend to make greater use of self-regulated learning strategies (Almayez et al., 2025; An et al., 2021; Bai & Wang, 2023), which in turn supports more productive strategy use and, ultimately, greater enjoyment. Conversely, enjoyment and intrinsic motivation may strengthen self-efficacy, particularly in technology-enriched environments, because learners ground their efficacy beliefs in experiences of success that autonomy helps to bring about (An et al., 2021; Dong et al., 2022a, b; Li & Wang, 2024; Mekheimer, 2025).

Enjoyment and motivation likely reinforce each other over time, but the present study positions motivation as an antecedent of enjoyment, following Self-Determination Theory (SDT) (Deci & Ryan, 2000). SDT holds that satisfying a learner's need for competence generates autonomous motivation, which in turn produces enjoyment. Control-Value Theory offers the same ordering from a different angle: value appraisals, which correspond to motivation, are theorized to precede and predict achievement emotions such as enjoyment (Pekrun, 2006; Pekrun & Perry, 2014). Two studies support this direction empirically: Liu et al. (2024) found that motivation predicts enjoyment in AI-mediated language learning, and Dewaele and Meftah (2024) found the same motivational antecedence in a pseudo-longitudinal design.

This same belief also shapes learners' willingness to take on demanding tasks, including EFL listening (Y. Wang & Liu, 2022). Motivation, framed through SDT, increasingly appears as the mechanism that carries this belief forward into affective outcomes such as enjoyment. Autonomous motivation in particular predicts engagement and enjoyment reliably across EFL settings (Alamer & Almulhim, 2021; Liu et al., 2025; Noels et al., 2014; Wang & Liu, 2022), while controlled motivation's predictive power depends heavily on context, and the relationships among the two motivation types, self-efficacy, and enjoyment often turn out nonlinear (Dong, Liu et al., 2022b). This complexity calls for models that can capture these parallel and interactive effects together, particularly in technology-mediated environments.

Longitudinal evidence has established that motivation and self-efficacy are strongly related and jointly account for a substantial proportion of the variance in EFL achievement (Fitriah et al., 2025; Mikami & Shiozawa, 2024). Their directionality, however, can vary by context: although motivation is typically treated as a precursor to self-efficacy, gains in enjoyment or self-efficacy can, in some cases, give rise to greater motivation (Li & Wang, 2024; Saito et al., 2025).

The three frameworks are used at different levels of explanation rather than as competing accounts. Social Cognitive Theory supplies the construct of self-efficacy and the claim that competence beliefs shape which activities learners approach and how much effort they invest. SDT supplies the account of motivation as varying in the degree to which it is self-endorsed. Control-Value Theory supplies the link between the two and the emotional outcome, holding that achievement emotions such as enjoyment arise from the joint appraisal of control over an activity and of its value. Motivation is positioned between self-efficacy and enjoyment on this basis, because a competence appraisal is theorized to inform the value a learner places on an activity, and enjoyment follows from both appraisals together. This ordering is theoretical, and the sensitivity analyses reported below show that the present cross-sectional data do not test it.

These environments also fit a broader shift in how the language classroom itself is organized: avatars increasingly function as interactional agents in their own right, and metaverse-based instruction can reach learners across institutions without the constraints of a physical classroom (Çelik & Yangın Ersanlı, 2026). Meta-analytic evidence indicates

that immersive and extended-reality environments produce a large overall effect on language learning outcomes (Chen et al., 2022), and metaverse-based listening tasks specifically have been shown to support the social-constructivist and collaborative skills that listening development depends on (Çelik & Yangın Ersanlı, 2025; Zhang, 2024). Such environments can also be designed deliberately for inclusion, for example by representing people with disabilities and other under-represented groups, which models socially responsible language use as part of the task itself. Watching or interacting with an avatar has separately been shown to increase learners' confidence and engagement, particularly for lower-achieving students, by lowering the affective filter typical of face-to-face classrooms (Han et al., 2023; Lan et al., 2018).

A central premise of the present study, and the dimension that most directly aligns it with research on personality, identity, and the self in digital environments, is the role of digital self-expression and avatar-based identity construction as enabling conditions for motivation and enjoyment in immersive learning environments. Research in digital psychology has established that individuals actively present, negotiate, and adapt their identities in online contexts through self-presentation strategies including profile design, communication style, and avatar customization. Experimental research has further shown that the characteristics of a digital avatar can shape users' own behavior and self-perception, a phenomenon termed the Proteus effect (Yee & Bailenson, 2007); a meta-analysis of 46 experimental studies confirmed a reliable, small-to-medium effect (Ratan et al., 2020). These processes extend classic accounts of self-presentation, in which individuals strategically manage the impressions they convey to others (Goffman, 1959), into digital environments where identity is enacted through avatars and mediated interaction. In language learning contexts specifically, avatar-mediated interaction has been shown to reduce anxiety and increase learner confidence, particularly among lower-achieving students who experience heightened evaluative threat in face-to-face settings (Han et al., 2023; Lan et al., 2018). The present study extends this line of inquiry by treating avatar personalization not merely as a technological feature of the Spatial.io platform but as a form of digital self-expression that learners may relate to their self-efficacy beliefs, their motivational orientations, and the quality of the listening experience. Because avatar personalization was common to every participant and was not measured, this relationship is examined only through participants' own retrospective accounts, and digital self-expression is treated as a qualitative feature of the experience rather than as an antecedent of the measured constructs. The present study focuses on this expressive, self-presentational dimension of avatar use rather than on how platform features shape that self-presentation or how traits might be inferred from behavioral data, which fall outside its single-session design.

Despite considerable evidence linking self-efficacy, motivation, and enjoyment, important gaps remain. Most existing models rely on cross-sectional data and therefore cannot support causal inference or capture how these constructs change dynamically over time (Zhang et al., 2024; Zhou & Wang, 2024). Immersive environments such as the metaverse are also notably underrepresented in this literature, which raises new questions about how interactivity, presence, and reduced social anxiety may reshape the previously established relationships among these constructs (Wang & MacIntyre, 2021; Wu & Kabilan, 2025).

Several recent studies have examined language learning in the metaverse, with populations, methods, and outcomes that vary considerably. Three used pre/post or comparative designs focused on instructional outcomes rather than on the psychological pathways examined here. Kuppusamy and Norman (2024) tracked perceptions of listening, self-confidence, and enjoyment in a training program for Malaysian secondary students using a pre-/post-test design; Zhang (2024) used structural equation modeling to test whether engagement or enjoyment mediates the effects of motivation and ambiguity tolerance on learning effectiveness; and J. Lee and Ko (2025) compared metaverse, face-to-face, and textbook-based instruction among adolescents, finding that the metaverse condition increased engagement and reduced peer pressure without outperforming traditional methods on listening itself. Across this work, motivation, interest, and self-efficacy consistently predict learning outcomes in metaverse-based EFL settings, but the samples are predominantly younger learners and none of these studies models how the constructs relate to one another. A separate, qualitative strand reaches a similar conclusion from different evidence: S.-M. Lee and Ahn (2025) found that a playful, constructivist metaverse fostered motivation, collaboration, and enjoyment among EFL learners, but, like the studies above, did not quantify the relationships among these constructs. The present study differs from all of this work in the same two respects: it uses university-level learners rather than adolescents, and it models self-efficacy, motivation, and enjoyment as an interrelated system, using path analysis and fsQCA, rather than treating them as separate outcomes or testing a specific intervention.

Taken together, these studies clarify the contribution of the present work. Research on metaverse-based language learning has established that immersive environments are associated with higher engagement and more positive affect, and research on foreign language enjoyment has established that self-efficacy and motivation are among its strongest correlates. These two literatures have developed largely in

parallel. No study has yet examined how self-efficacy, motivation, and enjoyment relate to one another inside an avatar-mediated metaverse listening environment, or how learners themselves account for the role of the avatar in that experience. The present study addresses this gap in three ways: it estimates the associations among the three constructs within a single immersive task; it uses a configurational analysis to ask which combinations of learner resources are sufficient for high or low enjoyment; and it uses learners' own accounts to describe what avatar-based digital self-expression meant to them. Avatar-based digital self-expression is treated throughout as an experiential feature of the task and as a qualitative construct, not as a measured antecedent.

RQ1. To what extent is self-efficacy associated with foreign language enjoyment directly and indirectly through motivation in an avatar-mediated metaverse listening environment?

RQ2. What configurational combinations of self-efficacy and motivation are sufficient for high or low foreign language enjoyment in this context?

RQ3. How do learners describe avatar-based digital self-expression in an avatar-mediated metaverse listening environment, and how do they relate it, in their own accounts, to their confidence, motivation, and enjoyment?

Method

Research design

This study used a convergent mixed-methods design. The quantitative and qualitative data were collected in a single administration immediately after the metaverse session, analyzed independently, and merged at the interpretation stage to produce meta-inferences. The quantitative strand estimated the associations among self-efficacy, motivation, and enjoyment; a fuzzy-set qualitative comparative analysis examined which combinations of self-efficacy and motivation were sufficient for high or low enjoyment; and the qualitative strand described how learners accounted for their own avatar experience. The three strands are integrated in the joint display presented in Table 4, which sets each quantitative finding alongside the configurational result and the qualitative codes that bear on it, and states whether the strands converge, diverge, or are silent on one another. The design is cross-sectional and single-group. There was no baseline measure, no comparison condition, and no random assignment, and exposure consisted of a single 10-minute session. All results are therefore reported as associations and as participants' retrospective perceptions. The study cannot establish that the metaverse environment or the avatars changed self-efficacy, motivation, or enjoyment.

The conceptual status of avatar-based digital self-expression

Avatar creation and personalization was an experiential feature of the task rather than a study variable. Every participant created an avatar, so the feature did not vary across participants; it was not manipulated, not assigned, and not measured with a quantitative instrument. Digital self-expression is therefore a qualitative construct in this study, derived from participants' written accounts of what creating and inhabiting an avatar was like for them. It is not an antecedent tested against self-efficacy, motivation, or enjoyment; its relationship to those constructs is examined solely through what participants reported. The quantitative model contains three variables: self-efficacy as the predictor, motivation as the intermediate variable, and enjoyment as the outcome.

Participants

A total of 131 undergraduate students were recruited from intact intermediate-level EFL classes at a public university in Türkiye. The EFL course is a compulsory, transversal English course required of all undergraduates regardless of discipline, meeting for four contact hours per week across the semester. The sample comprised 67 male (51.1%) and 64 female (48.9%) students aged 18 to 22 years (Table S1). All students were in the same proficiency stratum (intermediate/B1), as determined by institutional placement procedures. With respect to prior experience with immersive environments, 98 participants (74.8%) reported previous metaverse or VR experience, whereas 33 (25.2%) reported none. Most participants reported frequent use of digital technologies for learning (always, 21.4%; often, 42.7%; sometimes, 33.6%; never, 2.3%). Learners without prior VR or metaverse experience (25.2%) may have approached the platform from a different psychological vantage point than more experienced peers; the potential influence of this familiarity is noted as a limitation, and future research should examine whether prior immersive-technology experience moderates the relationships identified here.

Inclusion criteria were (a) enrollment in the targeted intermediate-level EFL course during data collection, (b) participation in the metaverse-based listening session on Spatial.io, (c) provision of written informed consent, and (d) completion of the post-session questionnaire. Exclusion criteria were (a) refusal or withdrawal of consent and (b) absence from the session or inability to access the platform. No participants met the exclusion criteria, and all completed questionnaires were retained. Because the study employed a single-condition, post-session design with intact classes, participants were not randomly assigned to

conditions, and all experienced the same metaverse-based listening task.

Sample size and statistical power

The sample comprised all students enrolled in the targeted intermediate-level EFL course during the data collection period, so its size was determined by the accessible population. A Monte Carlo sensitivity analysis was therefore conducted to establish what this design could detect. Five hundred datasets were simulated at $n = 131$ for each of several effect sizes, and the indirect association in each was tested with 600 bootstrap resamples. At $n = 131$, the bootstrap test had power of 0.94 to detect an indirect association built from standardized paths of a = b = 0.30, 0.78 for a = b = 0.26, and 0.48 for a = b = 0.20. The study was therefore well powered for the medium-sized indirect association observed here (a = 0.53, b = 0.47, power > 0.99) but underpowered for small ones. Precision is accordingly reported throughout as bootstrap confidence intervals. For the configurational analysis, 131 cases distributed across only four logically possible configurations yield 53, 17, 18, and 43 cases with membership above 0.5, so every configuration is empirically populated and none is a logical remainder.

Data collection tools

Four instruments provided the data analyzed in this paper: a demographic information form and three Likert-type scales, all administered in Turkish to ensure comprehension. The Foreign Language Learning Self-Efficacy Belief Scale (Adıyaman, 2019) was developed and validated in Turkish, so no translation was required. The English Language Learner Motivation Scale (Ardasheva et al., 2012) and the Foreign Language Enjoyment Scale (Dewaele & MacIntyre, 2014), originally in English, were translated into Turkish through a forward–backward procedure: a bilingual applied-linguistics expert translated the scales into Turkish, a second independent bilingual expert back-translated them into English, and discrepancies were resolved through discussion until consensus.

All items of all three scales were presented with a contextualizing stem directing participants to the session they had just completed ("In the metaverse, …"). This adaptation was made because the study asked about self-efficacy, motivation, and enjoyment as experienced during the avatar-mediated listening task rather than about foreign language learning in general, and because a general referent would have invited participants to report habitual beliefs formed in ordinary classrooms. The adaptation changes the referent of the items, and its implications for comparability with the published instruments are considered among the limitations. Reliability and validity evidence for the adapted, contextualized versions in the present sample is reported in Table 1.

Demographic information form

This form collected age, gender, and educational status using categorical response options and was administered first to contextualize responses to the subsequent scales.

Self-efficacy scale

Self-efficacy was measured with the Foreign Language Learning Self-Efficacy Belief Scale (Adıyaman, 2019). The published instrument contains 25 items loading on three factors, listening comprehension, written and oral expression, and reading comprehension, which together explain 52.8% of the variance, and it uses a 5-point Likert format (1 = strongly disagree, 5 = strongly agree). Adıyaman (2019) reported an overall Cronbach's alpha of 0.977, with test–retest coefficients of 0.914 for listening comprehension, 0.896 for written and oral expression, and 0.935 for reading comprehension. Fifteen of the 25 items were administered in the present study: the complete eight-item listening comprehension subscale together with seven items from the written and oral expression subscale, comprising five speaking items and two basic writing items. The ten remaining items, which concern extended writing and reading comprehension, were not administered because the metaverse task contained no reading passage (other than speech bubbles which transcribed what was being told) and no writing component, so those items had no referent in the activity

Table 1 Descriptive statistics, reliability, and convergent validity of the scales ($n = 131$)

Scale	Items	M	SD	Min	Max	α [95% CI]	ω	AVE	CR	√AVE
Self-efficacy	15	3.85	0.66	2.00	5.00	0.94 [0.93, 0.96]	0.95	0.57	0.95	0.75
Motivation (total)	12	3.98	0.75	1.17	5.00	0.92 [0.90, 0.94]	0.93	0.53	0.93	0.73
Intrinsic motivation	6	4.01	0.79	1.00	5.00	0.89 [0.86, 0.92]	0.90	0.61	0.90	—
Introjected regulation	2	3.72	1.08	1.00	5.00	0.90 [0.86, 0.93]	0.90	0.82	0.90	—
External regulation	4	4.07	0.81	1.00	5.00	0.81 [0.75, 0.86]	0.83	0.55	0.83	—
Enjoyment	21	4.02	0.79	1.95	5.00	0.97 [0.96, 0.98]	0.97	0.63	0.97	0.80

Note. The possible range of every composite is 1 to 5. AVE = average variance extracted; CR = composite reliability. Subscale values are reported for the sensitivity analysis in Table S3 and are not used in the main analyses

participants were asked to report on, and answering them would have required the general judgments the contextualizing stem was designed to exclude. A sample item is "I can extract the main idea of texts I hear in a foreign language." The measure is accordingly an adapted, task-relevant index of foreign language self-efficacy rather than the full published scale. In the present sample, α = 0.94, 95% CI [0.93, 0.96], and McDonald's ω = 0.95.

Motivation scale

Motivation was measured with the English Language Learner Motivation Scale (ELLMS): Pre-College (Ardasheva et al., 2012), which was developed from Noels et al.'s (2000) Language Learning Orientations Scale and is grounded in Self-Determination Theory. The instrument contains 12 items forming three subscales, intrinsic motivation (six items), introjected regulation (two items), and external regulation (four items), answered on a 5-point scale (1 = strongly disagree, 5 = strongly agree). Ardasheva et al. reported an overall Cronbach's alpha of 0.80 for the total validation sample (N = 1,057), with subscale alphas ranging from 0.58 to 0.74. All 12 items were administered and none was removed. Two items were adapted for the Turkish EFL setting: "I like to learn about Americans and how they live" became "I like to learn about Americans and English people and how they live," and "my American friends" became "my American and English friends." These changes were made because the participants learn English as a foreign language in Türkiye and have no systematic contact with American speakers, so the original wording would have had no referent for most respondents and would have measured contact opportunity rather than motivational orientation. A sample item is "It is fun to learn a new language." In the present sample, α = 0.92, 95% CI [0.90, 0.94], and ω = 0.93; subscale alphas were 0.89 for intrinsic motivation, 0.90 for introjected regulation, and 0.81 for external regulation.

Foreign language enjoyment scale

Enjoyment was measured with the 21-item Foreign Language Enjoyment Scale (Dewaele & MacIntyre, 2014), answered on a 5-point Likert format, for which the original authors reported α = 0.86. All 21 items were administered and none was removed. A sample item is "I enjoy it." Because the original items refer to the foreign language class, the teacher, and classmates, participants were instructed to answer them with reference to the metaverse session and to the avatars and peers present in it; without this instruction the items would have elicited judgments about the participants' regular classroom rather than about the task under study. In the present sample, α = 0.97, 95% CI [0.96, 0.98], and ω = 0.97.

Scoring

For each of the three scales a composite score was computed as the mean of the participant's item responses, so every composite has a possible range of 1 to 5, with higher scores indicating higher self-efficacy, motivation, or enjoyment. None of the items in the three scales is negatively worded, so no reverse coding was required. There were no missing responses: the online form required an answer to every item, and all 131 cases were complete on all 48 analyzed items.

Measurement quality in the present sample

Because two of the three instruments were translated, one was administered as a subset, and all three were contextualized to the metaverse session, evidence that the intended measurement structure was retained was examined in the present data. A sample of 131 is too small for a stable confirmatory factor analysis of 48 individual items, so measurement quality was assessed through internal consistency (Cronbach's α with its 95% confidence interval and McDonald's ω), through average variance extracted (AVE) and composite reliability (CR) computed from a single-factor solution for each scale, and through two discriminant validity criteria. Table 1 reports the results.

All three scales showed AVE above 0.50 and CR above 0.90, supporting convergent validity. Discriminant validity was assessed with the Fornell–Larcker criterion and the heterotrait–monotrait (HTMT) ratio of correlations. The square root of AVE exceeded every corresponding correlation for all three constructs: 0.75 for self-efficacy against correlations of 0.53 and 0.71, 0.73 for motivation against 0.53 and 0.71, and 0.80 for enjoyment against 0.71 and 0.71. HTMT ratios were 0.58 for self-efficacy with motivation, 0.75 for self-efficacy with enjoyment, and 0.76 for motivation with enjoyment, all below the conservative threshold of 0.85. The three constructs are therefore empirically distinguishable in these data.

Two qualifications remain. The correlations of enjoyment with self-efficacy and with motivation are both approximately 0.71, and because every measure was self-reported in a single administration, shared method variance is a plausible partial explanation for their magnitude and for the proportion of variance explained in enjoyment. In addition, the evidence reported concerns internal consistency and construct separation. A confirmatory factor analysis in a substantially larger sample is still required before the Turkish, session-contextualized versions of these instruments

can be treated as fully validated, and this is noted among the limitations.

Open-ended qualitative prompts

After the scales, participants responded in writing to three prompts capturing their experience of avatar-based digital identity expression and its perceived relationship to affective outcomes: (a) how designing and inhabiting their avatar affected their experience; (b) how they felt about expressing themselves through the avatar, and whether this influenced confidence or motivation; and (c) what connections they noticed between their avatar, their enjoyment, and their belief in their own ability. Prompts were administered in Turkish to encourage uninhibited expression and translated into English by the researcher for analysis; quotations are reproduced verbatim in translation. These prompts are directive: they explicitly invited participants to connect their avatar experience with confidence, motivation, and enjoyment. Any convergence between the qualitative accounts and the quantitative constructs is therefore partly built into the data collection procedure, and the qualitative strand is accordingly presented throughout as complementary elaboration of participants' perceptions.

Procedure

Data collection followed three stages: recruitment and briefing, the metaverse listening session, and survey administration. Course instructors announced the study and shared an information sheet describing the aims, the voluntary nature of participation, and assurances of anonymity and confidentiality. Students who agreed provided informed consent before data collection. No grades or credit were given, and students could withdraw at any time without penalty.

The listening experience took place in a public metaverse environment on Spatial.io, designed as an inclusive café that foregrounded the professional inclusion of individuals with mental and cognitive disabilities (Fig. 1). Within this environment, students encountered multiple animated avatars engaged in short, scripted English conversations about workplace interactions, customer service, and collaboration. All avatar speech was presented simultaneously as audio and as on-screen speech bubbles, supporting both auditory and visual processing. Dialogues were comprehensible at the intermediate (B1) level.

The role of artificial intelligence in the task

The avatars were animated, synthetically voiced non-player agents that delivered pre-scripted English dialogue on a loop as learners approached them. They did not generate language, did not adapt to what individual learners said or did, did not respond to learner input, and made no pedagogical decisions. Their contribution was to render fixed conversational input in an embodied, spatially situated, and multimodal form: learners chose which conversations to approach and heard each one as speech from a visible speaker while reading the same words in an on-screen bubble. The environment is accordingly characterized as avatar-mediated, and the activity as multimodal listening, since participants received the same content through two channels at once.

Pilot testing

Before the main data collection, and after ethical approval had been granted and all participants provided informed consent, the environment was piloted with a small group of students ($N=10$) to check audio quality, navigation, clarity of the instructions, and the difficulty of the dialogues. Minor adjustments to the instructions and to the positioning of the avatars followed. Pilot participants were not included in the main analytical sample of 131.

Sessions were conducted during regular class hours. The researcher introduced the environment, demonstrated basic navigation, and instructed students to attend to the English dialogues. Participants logged in on their own internet-connected devices, used headphones, remained in the same virtual café instance, and were given 10 min to explore, approach avatars, and listen to or read the scripted exchanges. No language test or performance evaluation was administered; the experience functioned as multimodal input exposure, and participants were subsequently asked to report their perceptions of self-efficacy, motivation, and enjoyment as experienced during it.

Immediately afterward, students exited the environment and completed the online questionnaire via Google Forms, individually and under supervision in a quiet classroom, taking approximately 15–20 min. Responses were then exported for analysis.

Qualitative data collection

Prior to entering the platform, each participant created and personalized a digital avatar using the platform's customization tools, selecting attributes such as appearance, clothing, and style. Consistent with self-presentation theory (Goffman, 1959), avatar personalization was treated as a form of digital self-presentation through which learners could negotiate and project an identity in the virtual environment. After the listening session and the questionnaire, participants responded in writing to the open-ended prompts. Quotations are identified by student number (S1–S131) to preserve anonymity.



Fig. 1 Avatar-mediated listening environment in the Spatial.io virtual café

Qualitative analytic approach

All 131 written responses were analyzed using qualitative content analysis. This choice follows from the analytic goal, which was to describe the distribution of accounts across the sample rather than to construct interpretive themes, and from the form of the evidence reported, namely how many participants expressed each idea. The analysis produced codes, that is, labeled units of meaning applied to whole responses. These codes were grouped into two categories by valence, positive and negative. The analysis stops at the level of codes and categories; no higher-order themes were developed.

Coding was primarily inductive. A first reading of all 131 responses generated an open set of candidate codes; these were consolidated into a codebook of eleven codes, each with a written definition and inclusion criteria; and all 131 responses were then recoded against the final codebook. Codes were applied at the level of the whole response, so a response could receive several codes, and the frequencies reported in Table S6 are counts of participants whose response contained evidence of the code.

A second coder, experienced in qualitative research in applied linguistics and not otherwise involved in the study, independently coded 26 of the 131 responses (19.8%), randomly selected from the full dataset, using the finalized codebook. Intercoder agreement was assessed before consensus using Cohen's kappa and indicated substantial agreement ($\kappa = 0.82$; code-level κs = 0.74–0.91). Disagreements were subsequently discussed until consensus was reached, and the consensus coding is reported in Table S6. The researcher facilitated the study sessions but was not involved in participants' formal assessment; to minimize potential researcher influence, responses were anonymized prior to analysis, coding decisions were guided by an explicit codebook, and a second independent coder was used to check the researcher's interpretations.

Data analysis

The indirect-association analysis was conducted in JASP (Version 0.95.4) using a regression-based path-analytic approach with maximum likelihood estimation, specifying self-efficacy as the predictor, enjoyment as the outcome, and motivation as the intermediate variable. Because the analysis was based on observed (manifest) variables rather than latent constructs, it constitutes regression-based path analysis rather than full structural equation modeling (SEM); accordingly, global SEM fit indices were not estimated, and model adequacy was evaluated through the direct and indirect paths and the proportion of explained variance. Indirect associations were estimated with bias-corrected bootstrapping using 5,000 resamples, and the bootstrap confidence interval, rather than the normal-theory z test, is treated throughout as the primary inferential evidence. All variables were standardized. Because self-efficacy, motivation, and enjoyment were measured in the same post-session administration, temporal ordering could not be established, and the indirect path is reported throughout as an indirect association consistent with the hypothesized model rather than as demonstrated mediation. Three alternative directional specifications were estimated as sensitivity analyses, and the assumptions of the fitted regression models were examined through residual diagnostics. To examine configurational pathways, a fuzzy-set qualitative comparative analysis (fsQCA) was conducted in fsQCA software (version 4.1) on calibrated composite scores for self-efficacy, motivation, and enjoyment, following the direct calibration method and the consistency and coverage formulas set out by Ragin (2008). Calibration, truth-table construction, and the consistency, proportional reduction in inconsistency (PRI), and coverage statistics were computed directly from the deposited composite scores, so that every value reported below is reproducible from the data and the anchors stated in the Findings. Supplementary Tables S1 to S6 are provided in the Supplementary Materials.

Findings

Preliminary analysis

Descriptive statistics for the three composites are reported in Table 1. Tests of univariate normality on observed composites bear only indirectly on a regression model, because the distributional assumption concerns the residuals rather than the variables, so the assumptions of the fitted models were examined directly. Linearity was supported: Ramsey's RESET test on the model predicting enjoyment from self-efficacy and motivation was not significant, $F(1, 127) = 1.33$, $p = 0.251$. Multicollinearity was not a concern, with variance inflation factors of 1.38 for both predictors. The residuals were close to symmetric (skewness = −0.07) with mild positive excess kurtosis (0.96); the Shapiro–Wilk test on the residuals was significant, $W = 0.970$, $p = 0.006$, which at $n = 131$ detects small departures. Homoscedasticity was not supported: the Breusch–Pagan test was significant, LM = 9.63, $p = 0.008$. Heteroscedasticity-consistent (HC3) standard errors are therefore reported alongside the conventional estimates, and the conclusions are unchanged: self-efficacy, β = 0.467, HC3 $SE = 0.085$, $p < 0.001$; motivation, β = 0.465, HC3 $SE = 0.099$, $p < 0.001$. Influence diagnostics identified no case with Cook's distance above 1 (maximum = 0.64) and eight cases above the 4/n convention of 0.031. Two cases had externally studentized residuals above 3 in absolute value (maximum = 3.40), and three cases exceeded the leverage cut-off of 3k/n. Refitting the model without the eight most influential cases left the pattern unchanged and slightly stronger (β = 0.486 for self-efficacy, β = 0.500 for motivation, $R^2 = 0.76$), so no case is driving the result, and all 131 cases were retained. Zero-order correlations among the composites were $r = 0.53$ for self-efficacy with motivation, 0.71 for self-efficacy with enjoyment, and 0.71 for motivation with enjoyment. Corrected for attenuation these become 0.56, 0.74, and 0.75, which remain below the level at which the constructs would be empirically indistinguishable, consistent with the discriminant validity evidence which was reported.

Direct, indirect, and total effects

Table 2 presents the estimates of the path model. The direct association between self-efficacy and enjoyment was significant (β = 0.467, $SE = 0.078$, $z = 5.97$, $p < 0.001$, 95% CI [0.306, 0.610]), indicating that higher self-efficacy was associated with greater enjoyment (RQ1).

The indirect association via motivation was also significant (β = 0.245, $SE = 0.064$, $z = 3.84$, $p < 0.001$). With the 5,000-resample bias-corrected bootstrap reported in the Method, the estimate was β = 0.245, bootstrap $SE = 0.066$, 95% CI [0.136, 0.385], and the percentile interval was [0.139, 0.390]; the indirect association accounts for 34.4% of the total association between self-efficacy and enjoyment. This is an indirect association consistent with the hypothesized model, for the reasons set out in the Method and examined further below (RQ1).

The total association was substantial and significant (β = 0.712, $SE = 0.048$, $z = 14.89$, $p < 0.001$, 95% CI [0.599, 0.789]). Because the model includes a single intermediate pathway, the total indirect association is numerically identical to the specific indirect association in Table 2.

Path coefficients and explained variance

Within the model, self-efficacy significantly predicted motivation (β = 0.526, $SE = 0.084$, $z = 6.24$, $p < 0.001$, 95% CI [0.320, 0.666]); motivation in turn significantly predicted enjoyment (β = 0.465, $SE = 0.068$, $z = 6.84$, $p < 0.001$, 95% CI [0.336, 0.604]); and the direct path from self-efficacy to enjoyment remained significant with motivation in the model (β = 0.467), which is the pattern the hypothesized model predicts (RQ1).

The note to Table 2 reports the explained variance for each endogenous variable, which describes explained variance and carries no information about model fit: the model accounted for 66.3% of the variance in enjoyment and 27.7% of the variance in motivation.

As summarized in Table 2, self-efficacy was directly associated with enjoyment (β = 0.47, $p < 0.001$) and also indirectly associated with enjoyment through motivation (β = 0.25, $p <$

Table 2 Path model estimates for self-efficacy, motivation, and enjoyment ($n = 131$)

Path	β	SE	z	*p*	95% CI
Self-efficacy → Motivation (a)	0.526	0.084	6.24	<0.001	[0.320, 0.666]
Motivation → Enjoyment (b)	0.465	0.068	6.84	<0.001	[0.336, 0.604]
Self-efficacy → Enjoyment, direct (c′)	0.467	0.078	5.97	<0.001	[0.306, 0.610]
Self-efficacy → Motivation → Enjoyment, indirect (a × b)	0.245	0.064	3.84	<0.001	[0.136, 0.385]
Self-efficacy → Enjoyment, total (c)	0.712	0.048	14.89	<0.001	[0.599, 0.789]

Standardized estimates from maximum likelihood estimation. CI = 95% bias-corrected bootstrap confidence interval, based on 5,000 resamples for the indirect association and 1,000 resamples for the other paths. The model explained 66.3% of the variance in enjoyment ($R^2 = 0.663$) and 27.7% of the variance in motivation ($R^2 = 0.277$)

0.001) (RQ1). Path coefficients are rounded to two decimals (self-efficacy→motivation = 0.53; motivation→enjoyment = 0.47; self-efficacy→enjoyment = 0.47). Residual terms (0.85 for motivation; 0.58 for enjoyment) were calculated as $\sqrt{(1-R^2)}$. The simultaneous presence of a direct association and an indirect association via motivation is the pattern that the hypothesized model implies.

Alternative directional models

Self-efficacy, motivation, and enjoyment are theorized to be reciprocally related, and all three were measured simultaneously, so three alternative directional specifications were estimated as sensitivity analyses, each with 5,000 bias-corrected bootstrap resamples (Table S2).

The results are unambiguous on one point. Model M2, in which motivation predicts self-efficacy which in turn predicts enjoyment, fits these data exactly as well as the hypothesized model M1: the same explained variance in enjoyment (R^2 = 0.66) and an indirect association of the same magnitude ($\beta = 0.246$ against $\beta = 0.245$). Model M3, in which enjoyment occupies the intermediate position, also produced an indirect estimate whose interval excluded zero. Only M4 produced a non-significant estimate. These cross-sectional data therefore cannot adjudicate between directional accounts. The ordering follows Social Cognitive Theory and Control-Value Theory, which place competence appraisals prior to value appraisals and both prior to achievement emotions; it is a theoretical commitment and is not established by the present design. All interpretation of the model below is qualified accordingly.

Motivation subdimensions

The ELLMS: Pre-College combines intrinsic and extrinsic orientations, so its total score represents global motivation rather than autonomous motivation specifically. The main analyses accordingly use the total score, and the construct is interpreted as global motivation. To establish whether the pattern depends on which orientation is examined, the model was re-estimated with each subscale in the intermediate position (Table S3).

The indirect association was of comparable magnitude when intrinsic motivation or external regulation stood in for the total score, and substantially smaller, though still distinguishable from zero, for introjected regulation. The pattern reported here is therefore a property of global motivation. Because the subscales are short, especially introjected regulation with two items, these analyses are reported as sensitivity checks and are not interpreted substantively.

Fuzzy-set qualitative comparative analysis

To examine which combinations of conditions are sufficient for high or low enjoyment, an fsQCA was conducted. The configurational analysis asks a different question from the regression model: the regression estimates how much, on average, each condition contributes net of the other, whereas the configurational analysis asks which combinations are sufficient for the outcome and allows the combination sufficient for high enjoyment to differ from the negation of the combination sufficient for low enjoyment. With two conditions the analysis is deliberately modest in scope, yielding four logically possible configurations, and its contribution is to test whether the two resources operate as substitutes or only in combination, and whether the path to low enjoyment mirrors the path to high enjoyment. It is not presented as independent confirmation of the regression result. Composite scores for self-efficacy (SE), motivation (MOT), and enjoyment (ENJ) were calibrated into fuzzy-set membership using Ragin's (2008) direct method, with the 5th, 50th, and 95th percentiles of the observed distribution as thresholds for full non-membership, the crossover point, and full membership (SE: 2.93, 3.867, 5.00; MOT: 2.88, 4.00, 5.00; ENJ: 2.81, 4.00, 5.00). Percentile anchors were used because the three constructs have no externally established substantive thresholds on a 1–5 Likert metric: there is no accepted value above which a learner counts in absolute terms as a high-enjoyment case, so membership was defined relative to the observed distribution. Because such anchors are sample-dependent, a robustness analysis under three alternative calibrations is reported below. Because the crossover anchors are sample medians, 18 participants obtained a membership score of exactly 0.5 in self-efficacy or motivation: 7 had a self-efficacy composite of 3.867, 13 had a motivation composite of 4.00, and 2 met both conditions. A case with a membership score of exactly 0.5 cannot be assigned to a truth-table row, so a constant of 0.001 was added to these scores (Fiss, 2011). The adjustment left every consistency, PRI, and coverage value unchanged at the third decimal and allowed all 131 cases to be assigned. With two conditions the truth table contains four rows, and all four are empirically populated: 53, 17, 18, and 43 cases respectively have membership above 0.5 in SE*MOT, SE*~MOT, ~SE*MOT, and ~SE*~MOT. A frequency threshold of one case therefore retains every row, there are no logical remainders, no counterfactual simplification was required, and the complex, intermediate, and parsimonious solutions are identical. A raw consistency threshold of 0.75 and a proportional reduction in inconsistency (PRI) threshold of 0.70 were applied for sufficiency, and 0.90 for necessity (Ragin, 2008).

Necessity analysis indicated that neither self-efficacy (consistency = 0.853) nor motivation (consistency = 0.856) met the 0.90 threshold, so neither is a necessary condition for high enjoyment; the negated conditions showed markedly lower consistency for that outcome (Table S4).

Sufficiency analysis is reported for the full truth table in Table 3, so that every configuration, including those that were not retained, is visible. Only one configuration met both the consistency and the PRI criteria for high enjoyment: the co-presence of high self-efficacy and high motivation (consistency = 0.939, PRI = 0.900, raw coverage = 0.692, 53 cases). The two single-resource configurations exceeded the raw consistency threshold of 0.75 but fell below the PRI threshold of 0.70 (PRI = 0.567 for SE*~MOT and 0.648 for ~SE*MOT). A low PRI indicates that these configurations are also substantially consistent with the negation of the outcome, so they cannot be treated as reliable pathways to high enjoyment, and their unique coverage was in any case small (0.065 and 0.056). The data therefore do not support the conclusion that either resource can substitute for the other. For low enjoyment, the simultaneous absence of both resources met both criteria under the primary calibration (consistency = 0.883, PRI = 0.806, raw coverage = 0.740, 43 cases). The absence of only one resource did not (PRI = 0.39 for SE*~MOT and 0.31 for ~SE*MOT). The two solutions are asymmetric. The logical negation of SE*MOT is ~SE + ~MOT, which would include every learner lacking either resource, whereas the solution for low enjoyment is limited to learners lacking both.

Because the calibration anchors are sample-dependent and the consistency threshold is a matter of analytic judgment, the solution was re-examined under three alternative calibrations and at raw consistency thresholds from 0.75 to 0.90 (Table S5).

Only one result holds under every specification in Table S5. SE*MOT met both criteria for high enjoyment under all four calibrations and at every raw consistency threshold from 0.75 to 0.90, with PRI between 0.886 and 0.973. The configuration ~SE*~MOT met both criteria for low enjoyment under the two percentile calibrations at thresholds up to 0.85. Its consistency stayed between 0.883 and 0.895 under every calibration, so it failed a threshold of 0.90, and its PRI fell to 0.595 and 0.673 under the substantive anchors. The single-resource configurations showed the reverse pattern. They failed the PRI criterion under both percentile calibrations and met both criteria under the substantive anchors at every threshold up to 0.90 (consistency = 0.912 to 0.938, PRI = 0.764 to 0.817). The substantive anchors place 117 of the 131 cases in the SE*MOT row, so each of the other three rows contains only 3 to 6 cases, and their statistics rest on very few learners. The sufficiency of SE*MOT for high enjoyment is therefore the robust configurational finding. The sufficiency of ~SE*~MOT for low enjoyment holds under distribution-based calibration only, and the sufficiency of either resource alone for high enjoyment remains undetermined. One further result was stable across all specifications: the absence of a single resource was never sufficient for low enjoyment (PRI ≤ 0.42).

Qualitative content analysis

Qualitative content analysis yielded eleven codes (seven positive, four negative); the five most prevalent positive codes formed the core of the account (Table S6). Avatar-based digital self-expression was the most prevalent code, present in 107 of the 131 responses (81.7%). Participants consistently described customizing their avatar as accompanying a psychologically distinct self-presentation, often more outgoing or confident than their offline self ("My real personality is very quiet but my avatar is very crazy and colorful. I think this digital me helped me for speak more," S3), consistent with accounts of digital self-presentation (Goffman, 1959). Foreign language enjoyment was also frequently articulated, often in contrast to prior classroom experience ("Honestly, this was the max fun I ever had in English lesson," S13).

Reported sense of competence was the second most prevalent code (77, 58.8%), with participants describing feeling more capable as the task went on ("I made mistakes but my confidence increased step by step. I am surprised," S7), consistent with Bandura's (1997) account of enactive mastery. Reported motivational shift (56, 42.7%) captured descriptions of a shift from extrinsic to intrinsic engagement ("My main motivation was just passing the course before this study. But in the café, I wanted to finish the game mission. I forgot about grades," S10). Reported sense of safety (37, 28.2%) captured descriptions of the avatar

Table 3 Analysis of sufficient conditions: full truth table (fsQCA)

Configuration	Outcome	Cons.	PRI	Raw cov.	Unique cov.	Cases (*n*)	Retained
SE * MOT	High ENJ	0.939	0.900	0.692	0.255	53	Yes
SE * ~MOT	High ENJ	0.820	0.567	0.396	0.065	17	No (PRI < 0.70)
~SE * MOT	High ENJ	0.844	0.648	0.451	0.056	18	No (PRI < 0.70)
~SE * ~MOT	Low ENJ	0.883	0.806	0.740	0.740	43	Yes

~ = absence of condition; * = logical AND. PRI = proportional reduction in inconsistency. Retained solution for high enjoyment (SE*MOT alone): consistency = 0.939, coverage = 0.692. Solution for low enjoyment (~SE*~MOT): consistency = 0.883, coverage = 0.740. Thresholds: raw consistency 0.75, PRI 0.70, frequency 1. All four truth-table rows are populated, so no logical remainders were simplified

as a psychological shield ("Behind my avatar, I felt safe. I am always scared of teacher's eyes in real class, but AI doesn't have eyes like that," S12). Two further positive codes engaged the value and autonomy dimensions of the theoretical frame: perceived authenticity (16, 12.2%) lent a sense of real-world language use ("The virtual café felt like I am studying in London or America. It was so real," S23), mapping onto the value dimension of Control-Value Theory, and autonomy and ownership (12, 9.2%) reflected control over movement and pace ("Spatial.io gave me freedom because I can move my avatar anywhere. I talked to AI and felt independent," S21), consistent with the Self-Determination Theory account in the hypothesized model.

Negative codes were markedly less prevalent. Task difficulty or comprehension gaps (30, 22.9%) typically did not occur in isolation, appearing alongside enjoyment or motivation in 22 of the 30 cases. Initial anxiety or apprehension (23, 17.6%) usually referenced pre-existing concerns about listening rather than the environment itself, and co-occurred with a reported sense of competence in 15 of the 23 cases. Skepticism or uncertainty (10, 7.6%) sometimes coincided with positive outcomes ("It is a strange connection but it worked for my motivation," S6), and lack of motivation was the least prevalent code (4, 3.1%). Because the least frequent codes were endorsed by relatively few participants (lack of motivation, $n=4$; skepticism, $n=10$), they are reported for transparency.

Integration of the three strands

Consistent with the convergent design described in the Method, the three strands were analyzed independently and merged at the interpretation stage. Table 4 presents the joint display, which sets each quantitative finding alongside the corresponding configurational result and the qualitative codes that bear on it, and states whether the strands converge, diverge, or are silent on one another. Because the open-ended prompts were directive, convergence between the qualitative and quantitative strands is treated as elaboration of participants' perceptions rather than as corroboration.

Discussion

Self-efficacy, motivation, and foreign language enjoyment (RQ1)

Students who saw themselves as more competent reported greater enjoyment ($\beta=0.467$, $p < 0.001$), consistent with Social Cognitive Theory (Bandura, 1997): efficacy beliefs are thought to foster positive affect by improving the balance learners perceive between a task's demands and their own skill. In the metaverse task, learners with higher self-efficacy appeared to treat interactions with the avatars as manageable rather than threatening, a pattern that accompanied more positive emotions during listening (Y. Wang, Derakhshan et al., 2021) — a dynamic that may matter even more in immersive environments, where the novelty of avatar-mediated communication could otherwise increase apprehension. This fits the broader positive-psychology turn in language education (Dewaele et al., 2019; Wu & Kabilan, 2025) and reports linking self-efficacy to motivation,

Table 4 Joint display: Integration of the quantitative, configurational, and qualitative strands

Focus	Quantitative strand	Configurational strand	Qualitative strand	Meta-inference
Self-efficacy and enjoyment	Direct association, $\beta=0.47$, 95% CI [0.31, 0.61]	SE alone fails PRI under percentile anchors (PRI = 0.57) and passes under substantive anchors (5 cases)	Reported sense of competence coded in 77 responses (58.8%)	Converge on association. The accounts describe competence gains, but no strand establishes direction.
Motivation as intermediate variable	Indirect association, $\beta=0.25$, 95% CI [0.14, 0.39]; alternative orderings fit equally well	MOT alone fails PRI under percentile anchors (PRI = 0.65) and passes under substantive anchors (6 cases)	Reported motivational shift coded in 56 responses (42.7%)	Converge on the involvement of motivation; diverge on whether it is intermediate, since the design cannot order the variables.
Combination of resources	Model including both predictors explained 66.3% of the variance in enjoyment	SE*MOT sufficient under all calibrations, PRI = 0.90, 53 cases	Co-occurrence of self-efficacy and motivation codes within responses	Converge: the two resources are informative together rather than separately.
Avatar-based digital self-expression	Not measured	Not modeled	Most prevalent code, 107 responses (81.7%)	Qualitative strand only. No quantitative claim is made, and the prompts invited this content.
Negative experience	Not modeled	~SE*~MOT sufficient for low enjoyment under percentile anchors only, 43 cases	Task difficulty (30), initial anxiety (23), skepticism (10), lack of motivation (4)	Converge tentatively: a substantial minority reported difficulty, which mean-level results obscure.

SE = self-efficacy, MOT = motivation, PRI = proportional reduction in inconsistency, CI = confidence interval, ~ = absence of condition, * = logical AND. Counts in the qualitative strand are numbers of participants whose responses contained the code ($N=131$)

enjoyment, and reduced burnout across EFL contexts (Song, 2024; Wu et al., 2022; Zhou & Wang, 2024). Practically, this suggests that learners' self-efficacy deserves deliberate attention in the design of avatar-mediated environments (Xu & Xu, 2025).

Self-Determination Theory (Deci & Ryan, 2000) predicts that this association should run partly through motivation, and the data are consistent with this prediction: the indirect association (β = 0.245, 95% CI [0.136, 0.385]) accounts for roughly 34% of the total association. This fits the idea that satisfying the need for competence supports self-endorsed motivation, which then sustains positive emotional engagement (Alamer et al., 2025; Ryan & Deci, 2020; Wang et al., 2024), and it matches other evidence that motivation links competence beliefs to positive affect in technology-enhanced learning (An et al., 2021). Two qualifications matter here. As Table S2 shows, a model in which motivation precedes self-efficacy reproduces the identical pattern, so the direct-plus-indirect result does not by itself prove that motivation mediates anything. And the motivation measure combines intrinsic and extrinsic orientations, so this finding describes global motivation rather than autonomous motivation specifically.

The model integrates Social Cognitive Theory with Control-Value Theory of achievement emotions (Pekrun, 2006, 2024). Self-efficacy functions as a control appraisal, learners' belief that they can perform the listening actions required to succeed, while motivation, particularly intrinsic motivation, represents the value dimension. The high explained variance ($R^2 = 0.66$) indicates that, together, these appraisals account for much of the variation in learners' enjoyment, consistent with evidence that control and value appraisals predict enjoyment both independently and interactively (Forsblom et al., 2022; Hawrot, 2024; Shao et al., 2020, 2026). Features of the task may bear on both appraisals. Multimodal delivery (audio plus on-screen text) offers learners multiple routes to comprehension, which earlier work has linked to lower anxiety than audio-only input (Islam & Wang, 2025; Lan et al., 2018; Tai & Chen, 2024), and inclusive, workplace-themed content may add to the perceived value and social relevance of the activity (Zhang, 2024). A comparison condition would be needed to test whether these features contributed to the appraisals reported here.

Configurational pathways to high and low enjoyment (RQ2)

The configurational analysis complements the symmetric model by shifting the question from the net contribution of each variable to the combinations of conditions that are sufficient for an outcome (Ragin, 2008; Sheng et al., 2025). Whereas the path analysis established that self-efficacy and motivation each contribute, on average, to enjoyment, the configurational analysis examines how these resources combine at the level of individual learners. One configuration was retained as sufficient for high enjoyment: the co-presence of high self-efficacy and high motivation (SE*MOT), which showed the highest consistency (0.939) and PRI (0.900) and covered roughly 69% of high-enjoyment cases. This conjunction aligns with both theoretical frames invoked here. In Control-Value Theory terms it represents the joint presence of a strong control appraisal and a strong value appraisal, the combination the theory predicts will yield the most robust achievement emotions (Pekrun, 2006; Shao et al., 2020); in Self-Determination Theory terms it reflects the co-occurrence of competence and the self-endorsed motivation that competence supports (Deci & Ryan, 2000; Ryan & Deci, 2020). That the configurational and regression analyses converge on the same conjunction lends coherence to the account, though the two draw on the same three variables and are not independent tests.

The single-resource configurations require closer scrutiny. Under the primary calibration, SE*~MOT and ~SE*MOT exceeded the raw consistency threshold but fell below the PRI threshold of 0.70, which means that each was also substantially consistent with the negation of the outcome, and both had small unique coverage. Under the substantive anchors, both met the consistency and PRI criteria at every threshold up to 0.90, yet each of these rows contained only five or six learners (Table S5). Whether a strong control appraisal or a strong value appraisal can each support enjoyment on its own therefore depends on how high and low membership are defined, and the present data leave the question open. It remains a hypothesis for samples with more learners at the low end of each scale.

The solution for low enjoyment was narrower than the negation of the solution for high enjoyment, which illustrates the set-theoretic asymmetry that distinguishes configurational from correlational analysis. Under the percentile calibrations, low enjoyment was associated with the joint absence of both resources (~SE*~MOT, raw coverage = 0.740, PRI = 0.806), and the absence of one resource alone was never sufficient for low enjoyment under any calibration. The sufficiency of the joint absence held only up to a consistency threshold of 0.85 and only when membership was defined relative to the sample distribution. Under the substantive anchors few learners scored low enough to belong to the low-enjoyment set, and the configuration fell below the PRI threshold. This result is accordingly treated as tentative. Descriptively, learners who brought at least one of the two resources to the task rarely reported low enjoyment. A comparison condition would be required to determine whether this pattern reflects anything about the environment itself.

The configurational findings carry a tentative design implication. The conjunction of both resources was the only configuration sufficient for high enjoyment under every specification, so instruction in comparable settings would be expected to need attention to self-efficacy and motivation together. Learners who arrive with neither resource may be the group for whom an immersive avatar-mediated task is least likely to be enjoyable, although this pattern held only under the percentile calibrations. If the pattern is confirmed, these learners are the group for whom additional scaffolding would matter most. More broadly, the regression model describes average associations across learners, and the configurational analysis describes which combinations of learner characteristics accompany high and low enjoyment.

Avatar-based digital self-expression in learners' accounts (RQ3)

Addressing RQ3, the analysis treats avatar-based digital self-expression as a concrete act of self-presentation through which learners construct and project a digital self, and asks how learners themselves relate this act to the constructs modeled in RQ1 and RQ2. Digital self-expression was not measured, not manipulated, and not entered into any model; the evidence consists entirely of what participants wrote after the session, in response to prompts that explicitly invited them to connect their avatar with their confidence, motivation, and enjoyment. Convergence between these accounts and the quantitative constructs is therefore partly built into the procedure and is reported as elaboration rather than as corroboration. Within those limits, the accounts were strikingly consistent. Participants described avatar creation as a form of self-presentation that they experienced as lowering evaluative exposure and as making the task feel manageable before it began; they described successful exchanges with the avatars as evidence of their own capability; and they described a shift in what they were working for during the session. The resemblance between these descriptions and the Proteus effect, in which users' self-perception aligns with the attributes of their avatar (Ratan et al., 2020; Yee & Bailenson, 2007), is notable. It is, however, a resemblance between the content of learners' retrospective descriptions and the content of a theory. The Proteus effect is accordingly offered as one plausible interpretation of what participants reported, alongside simple novelty and the reduced evaluative exposure of not being physically visible, and as a motivation for experimental work in which avatar attributes are manipulated directly.

Limitations

Several limitations should be noted. The sample was drawn from a single Turkish public university and included only intermediate-level students, constraining generalizability across proficiency levels, institutions, and sectors. The design is cross-sectional and single-group, with a single 10-minute exposure, no baseline measure, and no comparison condition, so the study describes associations and participants' retrospective perceptions and cannot establish that the environment or the avatars changed anything. Motivation was measured as a global construct, so the findings concern global motivation rather than autonomous motivation specifically, and no direct measure of listening performance was included, restricting conclusions about comprehension outcomes. The study was confined to a single platform (Spatial.io) and a single skill (listening), and it did not address the long-term scalability and equity of metaverse-based instruction across diverse populations and institutional contexts.

Several further limitations qualify these conclusions. The directional ordering of the variables is theoretical: Table S2 shows that a specification in which motivation precedes self-efficacy fits these data exactly as well as the one reported, with the same explained variance and an indirect association of the same magnitude, and cross-sectional estimates of indirect effects may in any case be biased (Maxwell & Cole, 2007). The reported direction should accordingly be read as a theoretically motivated interpretation, not as demonstrated mediation. The session itself was brief and novel: a single 10-minute immersive exposure raises the possibility that the affective responses partly reflect novelty rather than durable engagement. And because every construct was measured by self-report within one administration, common-method variance remains a plausible partial explanation for the size of the associations, including the 66% of variance explained in enjoyment.

The measurement itself carries limitations too. All three instruments were adapted for this study: every item carried a stem directing participants to the metaverse session, only 15 of the 25 self-efficacy items were administered, and two motivation items were reworded for the Turkish EFL context. These adaptations were necessary so that the measures referred to the activity participants were actually reporting on, and they are justified in the Method, but they mean that the psychometric properties established for the published instruments do not transfer automatically; the reliability and validity evidence in Table 1 supports the adapted measures without substituting for a confirmatory factor analysis in a larger sample. The three constructs were also strongly inter-correlated (rs = 0.53 to 0.71), and although the Fornell–Larcker criterion and the HTMT ratios supported discriminant validity, correlations this high do not fully rule out some conceptual overlap.

A further set of limitations concerns scope. The open-ended prompts asked participants directly about links between their avatar, their confidence, and their enjoyment, so the

qualitative material cannot serve as independent corroboration of the quantitative pattern. A quarter of the sample had no prior metaverse or virtual-reality experience, and the study could not test whether that prior experience moderates the relationships observed. The sample was adequately powered for medium-sized indirect associations but not for small ones, and the configurational analysis used only two conditions, which limits what it can add beyond the regression model. Finally, the study addresses only the expressive dimension of avatar-based identity; a fuller account would need designs that examine how digital self-presentation develops, is shaped by platform features, and is interpreted over longer periods of use.

Future research should use larger, multi-site samples; analyze motivational orientations separately with full-length subscale measures; test the directional ordering with longitudinal or experimental designs capable of establishing temporal precedence; incorporate behavioral listening assessments; and employ randomized or crossover designs (e.g., metaverse vs. non-metaverse, avatar vs. non-avatar) with baseline measures and delayed post-tests to strengthen causal inference and gauge retention. Cross-cultural replication and the use of structured digital-identity measures would further strengthen this account.

Conclusion

This study examined how self-efficacy, motivation, and foreign language enjoyment relate to one another in an avatar-mediated metaverse listening environment, and how learners describe avatar-based digital self-expression within it. Among 131 Turkish intermediate-level EFL undergraduates, self-efficacy was positively associated with enjoyment both directly and indirectly through motivation, and the two predictors together accounted for 66% of the variance in enjoyment. The sensitivity analyses show that these data are equally consistent with other directional orderings, so the indirect path is interpreted as an association. The configurational analysis identified the co-presence of high self-efficacy and high motivation as the only combination sufficient for high enjoyment under every calibration and threshold examined, and the joint absence of both as sufficient for low enjoyment under distribution-based calibrations only. Learners' written accounts added a further layer: they described creating and inhabiting an avatar as reducing evaluative exposure and making the task feel manageable, and they connected this directly to their confidence, motivation, and enjoyment. Taken together, the study supports an account in which a control-related belief and a value-related process accompany achievement emotions in immersive listening contexts, and it generates testable hypotheses about the role of digital self-presentation in that relationship.

Pedagogically, the findings suggest that avatar-mediated listening is more enjoyable for learners who bring both listening self-efficacy and motivation to the task, and that learners with neither may be the group least likely to find it enjoyable, although this pattern held only under distribution-based calibrations. For comparable Turkish university cohorts, instruction can accordingly combine graduated exposure to avatar conversations, scaffolded tasks that generate mastery experiences, and attention to the perceived relevance of the content, rather than targeting one resource alone. Given the brief, single-session, single-group design, these implications are best regarded as provisional and as hypotheses for design-based research rather than as established prescriptions. Instructors should also balance avatar-delivered input with human pedagogical guidance, normalize temporary failure as part of learning to listen, and treat avatar personalization as a pedagogical resource rather than an incidental feature. Future longitudinal work should examine how self-efficacy, motivation, and enjoyment co-develop during sustained metaverse use and whether avatar characteristics moderate the processes described by participants.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s12144-026-10370-8.

Acknowledgements N/A.

Author Contributions E.G. conceived and designed the study, collected and analyzed the data, and wrote the manuscript.

Funding This work received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data availability The data supporting the findings of this study are available from the author upon reasonable request.

Declarations

Ethics approval The study was approved by the Ethics Committee of Ondokuz Mayıs University (protocol code 2025-195; 28 February 2025) and was conducted in accordance with the Declaration of Helsinki.

Informed consent Informed consent was obtained from all participants prior to data collection.

Consent to publish All participants consented to the publication of their anonymized data and verbatim quotations. As all participants were adults (aged 18–22), no parental or guardian consent was required.

Competing interests The author declares no competing financial or non-financial interests that are directly or indirectly related to the work submitted for publication.

References

Adıyaman, A. (2019). Yabancı dil öğrenmeye yönelik özyeterlik inanç ölçeği: Geçerlik ve güvenirlik çalışması. *International Journal of Language Academy*, *7*(28), 80–92. https://doi.org/10.18033/ijla.4128

Alamer, A., & Almulhim, F. (2021). The interrelation between language anxiety and self-determined motivation: A mixed methods approach. *Frontiers in Education*, *6*. https://doi.org/10.3389/feduc.2021.618655. Article 618655.

Alamer, A., Robat, E. S., Shirvan, M. E., & Ryan, R. (2025). Self-determination theory and language learning: A multilevel meta-analysis. *Educational Psychology Review*, *37*(2). https://doi.org/10.1007/s10648-025-10038-y

Almayez, M., Al-Khresheh, M., Al-Qadri, A., Alkhateeb, I., & Alomaim, T. (2025). Motivation and English self-efficacy in online learning applications among Saudi EFL learners: Exploring the mediating role of self-regulated learning strategies. *Acta Psychologica*, *254*, 104796. https://doi.org/10.1016/j.actpsy.2025.104796

An, Z., Wang, C., Li, S., Gan, Z., & Li, H. (2021). Technology-assisted self-regulated English language learning: Associations with English language self-efficacy, English enjoyment, and learning outcomes. *Frontiers in Psychology*, *11*, 558466. https://doi.org/10.3389/fpsyg.2020.558466

Ardasheva, Y., Tong, S. S., & Tretter, T. R. (2012). Validating the English Language Learner Motivation Scale (ELLMS): Pre-college to measure language learning motivational orientations among young ELLs. *Learning and Individual Differences*, *22*(4), 473–483. https://doi.org/10.1016/j.lindif.2012.03.001

Bai, B., & Wang, J. (2023). The role of growth mindset, self-efficacy and intrinsic value in self-regulated learning and English language learning achievements. *Language Teaching Research*, *27*(1), 207–228. https://doi.org/10.1177/1362168820933190

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.

Çelik, F., & Yangın Ersanlı, C. (Eds.). (2026). *Metaverse-based language learning and teaching: Innovative, immersive approaches and future directions* (Vol. 49). Springer. https://doi.org/10.1007/978-3-032-31366-9

Çelik, F., & Yangın Ersanlı, C. (2025). Developing and integrating metaverse into EFL listening skills to increase learners' self-determination based on social constructivist theory. *Journal of Computer Assisted Learning*, *41*(6). https://doi.org/10.1111/jcal.70151. Article e70151.

Chen, J., Dai, J., Zhu, K., & Xu, L. (2022). Effects of extended reality on language learning: A meta-analysis. *Frontiers in Psychology*, *13*, 1016519. https://doi.org/10.3389/fpsyg.2022.1016519

Deci, E. L., & Ryan, R. M. (2000). The what and why of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01

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

Dewaele, J. M., & Meftah, R. (2024). How motivation and enjoyment spread their wings during English foreign language learning: A pseudo-longitudinal investigation into Moroccan learners. *Learning and Motivation*, *85*, 101962. https://doi.org/10.1016/j.lmot.2024.101962

Dewaele, J. M., Chen, X., Padilla, A. M., & Lake, J. (2019). The flowering of positive psychology in foreign language teaching and acquisition research. *Frontiers in Psychology*, *10*, 2128. https://doi.org/10.3389/fpsyg.2019.02128

Dong, L., Jamal Mohammed, S., Ahmed Abdel-Al Ibrahim, K., & Rezai, A. (2022a). Fostering EFL learners' motivation, anxiety, and self-efficacy through computer-assisted language learning-and mobile-assisted language learning-based instructions. *Frontiers in Psychology*, *13*, 899557. https://doi.org/10.3389/fpsyg.2022.899557

Dong, L., Liu, M., & Yang, F. (2022b). The relationship between foreign language classroom anxiety, enjoyment, and expectancy-value motivation and their predictive effects on Chinese high school students' self-rated foreign language proficiency. *Frontiers in Psychology*, *13*, 860603. https://doi.org/10.3389/fpsyg.2022.860603

Fiss, P. C. (2011). Building better causal theories: A fuzzy set approach to typologies in organization research. *Academy of Management Journal*, *54*(2), 393–420. https://doi.org/10.5465/amj.2011.60263120

Fitriah, L. M., Mujiono, M., & Weganofa, R. (2025). The interplay of learning motivation and self-efficacy on English learning outcomes of EFL learners: A meta-analysis. *IDEAS: Journal on English Language Teaching and Learning Linguistics and Literature*, *12*(2), 2756–2776. https://doi.org/10.24256/ideas.v12i2.5849

Forsblom, L., Pekrun, R., Loderer, K., & Peixoto, F. (2022). Cognitive appraisals, achievement emotions, and students' math achievement: A longitudinal analysis. *Journal of Educational Psychology*, *114*(2), 346–367. https://doi.org/10.1037/edu0000671

Goffman, E. (1959). *The presentation of self in everyday life*. Doubleday.

Han, E., Miller, M.R., DeVeaux, C., Jun, H., Nowak, K.L., Hancock, J.T., Ram, N., & Bailenson, J.N. (2023). People, places, and time: A large-scale, longitudinal study of transformed avatars and environmental context in group interaction in the metaverse. *Journal of Computer-Mediated Communication*, 28(2). https://doi.org/10.1093/jcmc/zmac031

Hawrot, A. (2024). Changes in control and value appraisals predict changes in learning enjoyment: A four-wave analysis among lower secondary school students. *British Journal of Educational Psychology*, *94*(1), 231–247. https://doi.org/10.1111/bjep.12644

Islam, M. Z., & Wang, G. (2025). Avatars in the educational metaverse. *Visual Computing for Industry Biomedicine and Art*, *8*, Article 15. https://doi.org/10.1186/s42492-025-00196-9

Kuppusamy, R., & Norman, H. (2024). Investigating the effectiveness of metaverse-based training for improving listening skills. *International Journal for Multidisciplinary Research*, *6*(3). https://doi.org/10.36948/ijfmr.2024.v06i03.22806

Lan, Y. J., Fang, W. C., Hsiao, I. Y. T., & Chen, N. S. (2018). Real body versus 3D avatar: The effects of different embodied learning types on EFL listening comprehension. *Educational Technology Research and Development*, *66*(3), 709–731. https://doi.org/10.1007/s11423-018-9569-y

Lee, S. M., & Ahn, T. Y. (2025). L2 learner experiences in a playful constructivist metaverse space. *ReCALL*, *37*(1), 129–145. https://doi.org/10.1017/S0958344024000235

Lee, J., & Ko, Y. (2025). Comparing metaverse and face-to-face instruction for enhancing English skills and attitudes among EFL adolescents. *Education and Information Technologies*, *30*(6), 8215–8244. https://doi.org/10.1007/s10639-024-13162-4

Li, J., & Wang, C. (2024). Measuring Chinese English as a foreign language learners' self-efficacy in vocabulary learning: Instrument development and validation. *Education Sciences*, *14*(6). https://doi.org/10.3390/educsci14060671

Liu, G. L., Darvin, R., & Ma, C. (2024). Unpacking the role of motivation and enjoyment in AI-mediated informal digital learning of English (AI-IDLE): A mixed-method investigation in the Chinese context. *Computers in Human Behavior*, *160*, 108362. https://doi.org/10.1016/j.chb.2024.108362

Liu, G. L., Zou, M. M., Soyoof, A., & Chiu, M. M. (2025). Untangling the relationship between AI-mediated informal digital learning of English (AI-IDLE), foreign language enjoyment and the ideal L2

self: Evidence from Chinese university EFL students. *European Journal of Education*, *60*(1). https://doi.org/10.1111/ejed.12846. Article e12846.

Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods*, *12*(1), 23–44. https://doi.org/10.1037/1082-989X.12.1.23

Mekheimer, M. (2025). Technological self-efficacy, motivation, and contextual factors in advanced EFL e-learning: A mixed-methods study of strategy use and satisfaction. *Humanities and Social Sciences Communications*, *12*, Article 677. https://doi.org/10.1057/s41599-025-04947-0

Mikami, H., & Shiozawa, T. (2024). A self-regulation perspective on L2 grit development and its impact on language achievement. *Language Teaching Research*. Advance online publication. https://doi.org/10.1177/13621688241287501

Noels, K. A., Pelletier, L. G., Clément, R., & Vallerand, R. J. (2000). Why are you learning a second language? Motivational orientations and self-determination theory. *Language Learning*, *50*(1), 57–85. https://doi.org/10.1111/0023-8333.00111

Noels, K. A., Chaffee, K. E., Michalyk, M., & Sugita McEown, M. (2014). Culture, autonomy, and the self in language learning. In K. Csizér & M. Magid (Eds.), *The impact of self-concept on language learning* (pp. 206–226). Multilingual Matters.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review*, *18*(4), 315–341. https://doi.org/10.1007/s10648-006-9029-9

Pekrun, R. (2024). Control-value theory: From achievement emotion to a general theory of human emotions. *Educational Psychology Review*, *36*(3). https://doi.org/10.1007/s10648-024-09909-7

Pekrun, R., & Perry, R. P. (2014). Control-value theory of achievement emotions. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 120–141). Routledge.

Ragin, C. C. (2008). *Redesigning social inquiry: Fuzzy sets and beyond*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226702797.001.0001

Ratan, R., Beyea, D., Li, B. J., & Graciano, L. (2020). Avatar characteristics induce users' behavioral conformity with small-to-medium effect sizes: A meta-analysis of the Proteus effect. *Media Psychology*, *23*(5), 651–675. https://doi.org/10.1080/15213269.2019.1623698

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, *61*, 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Saito, K., Dewaele, J. M., In'nami, Y., & Abe, M. (2025). Disentangling the causal role of motivation, enjoyment, and anxiety in second language speech learning: A final report. *Studies in Second Language Acquisition*, *47*(2), 461–487. https://doi.org/10.1017/S0272263125000038

Shao, K., Pekrun, R., Marsh, H. W., & Loderer, K. (2020). Control-value appraisals, achievement emotions, and foreign language performance: A latent interaction analysis. *Learning and Instruction*, *69*, 101356. https://doi.org/10.1016/j.learninstruc.2020.101356

Shao, K., Stockinger, K., Marsh, H. W., & Pekrun, R. (2026). Applying control-value theory for examining multiple emotions in L2 classrooms: Validating the Achievement Emotions Questionnaire-Second Language Learning. *Language Teaching Research*, *30*(1), 288–316. https://doi.org/10.1177/13621688221144497

Sheng, D., Pu, L., & Liu, H. (2025). The predictive effects of burnout, academic buoyancy and enjoyment on students' English academic achievement: A fsQCA approach. *Behavioral Sciences*, *15*(11). https://doi.org/10.3390/bs15111471

Song, Y. (2024). Assessing the interactions between learning enjoyment, motivation, burnout, and grit in EFL students: A mixed-methods approach. *BMC Psychology*, *12*, Article 796. https://doi.org/10.1186/s40359-024-02303-6

Tai, T. Y., & Chen, H. H. J. (2024). The impact of intelligent personal assistants on adolescent EFL learners' listening comprehension. *Computer Assisted Language Learning*, *37*(3), 433–460. https://doi.org/10.1080/09588221.2022.2040536

Wang, Y., & Liu, H. (2022). The mediating roles of buoyancy and boredom in the relationship between autonomous motivation and engagement among Chinese senior high school EFL learners. *Frontiers in Psychology*, *13*, 992279. https://doi.org/10.3389/fpsyg.2022.992279

Wang, L., & MacIntyre, P. (2021). Second language listening comprehension: The role of anxiety and enjoyment in listening metacognitive awareness. *Studies in Second Language Learning and Teaching*, *11*(4), 491–515. https://doi.org/10.14746/sslt.2021.11.4.2

Wang, Y., Derakhshan, A., & Zhang, L. J. (2021). Researching and practicing positive psychology in second/foreign language learning and teaching: The past, current status and future directions. *Frontiers in Psychology*, *12*, 731721. https://doi.org/10.3389/fpsyg.2021.731721

Wang, Y., Shen, B., & Yu, X. (2021). A latent profile analysis of EFL learners' self-efficacy: Associations with academic emotions and language proficiency. *System*, *103*. https://doi.org/10.1016/j.system.2021.102633. Article 102633.

Wang, Y., Wang, H., Wang, S., Wind, S. A., & Gill, C. (2024). A systematic review and meta-analysis of self-determination-theory-based interventions in the education context. *Learning and Motivation*, *87*, 102015. https://doi.org/10.1016/j.lmot.2024.102015

Wu, W., & Kabilan, M. K. (2025). Foreign language enjoyment in language learning from a positive psychology perspective: A scoping review. *Frontiers in Psychology*, *16*, 1545114. https://doi.org/10.3389/fpsyg.2025.1545114

Wu, X., Yang, H., Liu, J., & Liu, Z. (2022). English use anxiety, motivation, self-efficacy, and their predictive effects on Chinese top university students' English achievements. *Frontiers in Psychology*, *13*, 953600. https://doi.org/10.3389/fpsyg.2022.953600

Xu, J., & Xu, Y. (2025). The impact of self-efficacy on psychological resilience in EFL learners: A serial mediation model. *BMC Psychology*, *13*, 858. https://doi.org/10.1186/s40359-025-03236-4

Yee, N., & Bailenson, J. (2007). The Proteus effect: The effect of transformed self-representation on behavior. *Human Communication Research*, *33*(3), 271–290. https://doi.org/10.1111/j.1468-2958.2007.00299.x

Zhang, Q. (2024). Mediation/moderation effects of engagement, foreign language enjoyment, and ambiguity tolerance in metaverse-based foreign language learning. *International Journal of Educational Technology in Higher Education*, *21*. https://doi.org/10.1186/s41239-024-00484-z. Article 54.

Zhang, X., Wang, J., & Ke, X. (2024). An investigation into academic burnout, enjoyment and engagement in EFL learning among Chinese junior high school students. *Frontiers in Psychology*, *14*, 1292772. https://doi.org/10.3389/fpsyg.2023.1292772

Zhou, M., & Wang, X. (2024). The influence of enjoyment, boredom, and burnout on EFL achievement: Based on latent moderated structural equation modeling. *PLOS ONE*, *19*(9). https://doi.org/10.1371/journal.pone.0310281. Article e0310281.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.