Accounting, Marketing, and Organization, 2025, 1(3), 33-52
https://doi.org/10.71204/wb393133

The Influence Mechanism of Language Arousal in AI Digital Human Teachers on Learning Experience: The Mediating Role of Psychological Distance

Zirui Wang [1,*]

[1] Hebei University, Baoding071002, China

*** Correspondence:**
Zirui Wang

15532180079@163.com

Received: 28 November 2025 / Accepted: 30 December 2025 / Published online: 31 December 2025

Abstract

While AI digital human teachers are increasingly utilized in the field of education, current designs often prioritize technical implementation while overlooking the psychological adaptability mechanisms inherent in instructional communication. (1) Grounded in social presence theory and cognitive load theory, this study investigates the impact of AI teachers' language arousal levels (high vs. low) on students' learning experiences and examines the pivotal mediating role of psychological distance in instructional outcomes. (2) A single-factor between-subjects experimental design was employed, where 100 participants were randomly assigned to either a high-arousal language group (enthusiastic and vivid) or a low-arousal language group (steady and calm) for a learning session. (3) Data analysis reveals that language arousal significantly and positively affects instructional outcomes ($p<0.001$); specifically, the high-arousal strategy yielded significantly higher scores in learning satisfaction, continuance intention, and perceived learning effectiveness compared to the low-arousal group. Furthermore, psychological distance serves as a key mediator, as high-arousal language conveys emotional cues that bridge the gap between students, the AI teacher, and the instructional content, thereby reducing cognitive defenses. (4) This study identifies psychological distance as a critical psychological mechanism for successful AI instruction, providing empirical evidence for the anthropomorphic design of AI educational tools and facilitating a paradigm shift in AI teachers from mere speech synthesis to emotional instructional communication.

Keywords: AI Digital Human Teacher; Language Arousal; Psychological Distance; Instructional Communication; Learning Effectiveness

1. Introduction

With the rapid iteration and upgrading of artificial intelligence technologies, generative AI and digital human technologies are reshaping service and interaction experiences at an unprecedented pace. Artificial intelligence is evolving from purely analytical tools into "feeling" and interactive entities possessed of highly anthropomorphic characteristics, a transformation that demonstrates immense potential for enhancing user engagement (Huang and Rust, 2018). This trend has profoundly impacted the field of education, where AI digital human teachers, as an emerging educational medium, are gradually becoming a vital supplement to online education due to their advantages of 24/7 availability and personalization. However, although current technologies can endow AI teachers with realistic facial expressions, students often still experience a sense of "technological alienation" when facing virtual instructors. This psychological barrier frequently restricts deep instructional interaction and emotional investment.

In the deep logic of human-computer interaction, language serves not only as a carrier for knowledge transmission but also as a core cue for emotional communication and the construction of social presence. Expressed arousal—defined as the emotional intensity and energy conveyed by a speaker through language—can significantly alter audience perception through non-linear paths (Yin et al., 2017). Transposing this perspective to educational contexts, we posit that the "language arousal level" of AI teachers (manifested in intonation variation, rhythmic vitality, and emotional engagement) may determine the "temperature" of the interaction more critically than mere "visual realism."

However, a review of current educational technology literature reveals that research predominantly focuses on the algorithmic accuracy or visual design of AI. There is a relative scarcity of exploration into the micro-mechanisms of "how AI teachers influence learning experiences through anthropomorphic language strategies." particularly in a context where visual images are becoming increasingly homogenized, voice characteristics are emerging as a key differentiating variable determining the quality of human-computer interaction. Although empirical studies indicate that enhancing storytelling richness in immersive virtual reality environments can significantly promote science learning outcomes, such research has yet to deeply reveal the psychological transmission mechanisms behind this enhancement (Zhang and Bowman, 2022). In other words, academia still needs to answer a critical question: Through what psychological path does high-arousal language transform "cold algorithms" into "warm mentors" in the perception of students?

Psychological Distance Theory offers a unique theoretical lens for addressing this question. Relevant studies point out that perceived psychological distance plays a crucial buffering role in alleviating negative emotions and enhancing a sense of control (Zheng et al., 2020). In human-AI instructional interactions, high-arousal, vivid language may function as a potent social cue, effectively shortening the "interpersonal distance" between students and virtual teachers, as well as the "cognitive distance" between students and unfamiliar knowledge. This reduction in psychological distance may subsequently lower students' cognitive defenses and promote sustained willingness to learn and knowledge internalization.

Addressing the aforementioned theoretical gaps, this study aims to construct a mediation model to explore the impact of AI digital human teachers' language arousal levels (high vs. low) on students' learning experiences (including satisfaction, continuance intention, and perceived learning effectiveness), with a specific focus on examining the core mediating role of psychological distance. This study not only responds to theoretical calls for deepening the mechanisms of AI emotional interaction (Huang and Rust, 2018) but also attempts to provide empirical evidence and design guidance for the iteration of AI educational products from "functional substitutes" to "emotional companions."

2. Literature Review

2.1. Language Expectancy Theory and the "Feeling" Evolution of AI Digital Human Teachers

With the development of artificial intelligence technology, the role of AI in the service sector is undergoing a profound paradigm shift. Artificial intelligence is evolving from "Mechanical" AI, which performs repetitive tasks, and "Analytical" AI, which processes data, toward "Feeling" AI capable of emotional interaction. This evolution requires AI digital human teachers not only to accurately transmit knowledge but also to establish effective emotional connections (Huang and Rust, 2018).

Language Expectancy Theory (LET) provides a theoretical framework for understanding this process. The theory posits that individuals hold established normative expectations regarding the language use of communicators in specific contexts, and these expectations directly influence the effectiveness of information reception (Burgoon et al., 2002). In educational scenarios, virtual avatars should adopt a high-arousal language style that aligns with user expectations to compensate for the lack of vitality inherent in their virtual identity, thereby producing a positive dramatic effect (Wang et al., 2025).

However, the impact of language arousal (or expressed arousal) on user perception is not a simple linear growth. Empirical studies have found a non-linear relationship between the emotional arousal of text and the perceived helpfulness by the audience: while moderate emotional expression can enhance the helpfulness of reviews, excessive emotional venting may be counterproductive (Yin et al., 2017). The same logic applies to the educational field, where enhancing storytelling richness—through vivid, high-arousal narrative styles—has been proven to significantly promote students' understanding of scientific content and learning outcomes (Zhang and Bowman, 2022).

Accordingly, this study argues that AI teachers need to find an "optimal arousal zone" that satisfies students' expectations for "Feeling" AI through vivid narrative styles while avoiding cognitive interference caused by excessive arousal. Thus, Hypothesis H1 is proposed: Language arousal level significantly influences instructional outcomes (including learning satisfaction, continuance intention, and perceived learning effectiveness).

2.2. The Mediating Role of Psychological Distance: Sense of Control and Cognitive Proximity

Psychological Distance Theory suggests that individuals construct cognitive representations of an object based on its psychological proximity (Liberman and Trope, 2014). In human-AI instructional interactions, students often perceive a greater psychological distance due to the "coldness" of technology, which constitutes a barrier to deep interaction.

Psychological distance affects not only intimacy but also serves as a crucial psychological buffering mechanism. Research indicates that psychological distance plays a key mediating role between perceived sense of control and mental health; closer psychological distance helps enhance an individual's sense of control and alleviate negative emotions (Zheng et al., 2020). Transposed to the educational context, when AI teachers shorten psychological distance through high-arousal language (such as enthusiastic intonation and first-person narratives), they are effectively reducing the anxiety and cognitive defense students face when encountering unfamiliar knowledge (Weidlich et al., 2024).

This "distance-shortening" effect is manifested at two levels: (1) Shortened psychological distance enhances perceived intimacy and trust, making students more willing to view the AI as a "partner" rather than a "tool" (Jones et al., 2017; Massara and Severino, 2013). (2) Closer psychological distance promotes deep information processing and absorption, thereby improving knowledge retention performance.

Thus, Hypothesis H2 is proposed: Psychological distance plays a mediating role between language arousal and instructional outcomes. That is, moderately high-arousal language can shorten psychological distance, thereby enhancing students' sense of control and intimacy, and ultimately improving the learning experience.

2.3. The Auxiliary Observational Perspective of Emotional Resonance and Attention

In addition to the core path of psychological distance, the degree of language arousal also directly affects the audience's emotional resonance and concentration. Emotions act as "memory anchors" in learning contexts, significantly enhancing memory encoding (Barclay, 2020). Learning processes accompanied by positive emotional experiences often result in better knowledge retention (Wang et al., 2023).

However, the relationship between attention and arousal level requires a dialectical view. Although high-arousal language can attract attention through novelty and rhythm (Berger et al., 2023), excessively high arousal may lead to cognitive overload, conversely interfering with the acquisition of core information (Ludwig et al., 2013). Therefore, we treat emotional resonance and attention as auxiliary variables to comprehensively examine the instructional effects of language arousal.

Thus, Hypothesis H3 is proposed: Emotional resonance and attention act as auxiliary observational indicators and play a role in the process by which language arousal affects instructional outcomes.

3. Method

3.1. Participants and Design

This study consisted of two phases: a pre-experiment and a formal experiment. The pre-experiment aimed to verify the validity of the experimental materials and recruited 80 college student participants from a university in China, with 40 assigned to the high-arousal group and 40 to the low-arousal group. The formal experiment employed a two-condition between-subjects design (language arousal: high vs. low). A total of 100 college students were recruited (53 *females*, 47 *males*; M_{age}= 21.56, SD = 1.73. All participants were randomly assigned to the experimental groups and had not previously participated in similar experiments. The study strictly adhered to academic research ethical guidelines during implementation. Participation was anonymous, personal privacy was strictly protected, and all collected data were used solely for academic analysis.

3.2. Procedure and Materials

The stimulus videos used for material development were created using the "Jimeng" 3D digital human generation platform to establish a digital human teacher image with a highly unified appearance, thereby ensuring visual consistency throughout the experiment. The script content was uniformly selected from Introduction to Management, ensuring complete consistency in information structure and knowledge point coverage between the high- and low-arousal groups. However, experimental manipulation of arousal levels was achieved through the precise regulation of language expression parameters: the script for the high-arousal group adopted an expressive linguistic style, with the digital teacher's intonation modulated, stress clearly accentuated, and filled with passion; conversely, the low-arousal group's script adopted a calm and steady narrative style with minimal intonation fluctuation. This resulted in two sets of experimental video materials, each 18 seconds in length. This duration was designed to simulate the "micro-learning" scenarios prevalent in the current mobile internet environment, focusing on examining students' instantaneous reception of knowledge points and the formation of emotional impressions during fragmented learning. All post-processing and synthesis of the videos were completed using the video editing software "Jianying" (CapCut). To maximize the elimination of interference from extraneous variables, the font size and color of subtitles, the position of the digital teacher on the screen, the amplitude of body movements, clothing styling, and audio timbre were kept strictly constant throughout the experiment.

The image of the digital human teacher was set as a young male wearing a white shirt and black trousers, with a standard classroom blackboard scene as the background, presenting a friendly and natural overall visual style. During the experiment, all videos were displayed on an LCD screen equipped with high-fidelity speakers. This equipment combination aimed to approximate the auditory quality and sound effects of a real digital human teacher's lecture as closely as possible, simulating an audiovisual experience near that of a real classroom.

The manipulation check results from the pre-experiment confirmed the success of the material development. The perceived language arousal score of the high-arousal group (M = 5.39, SD = 1.00) was significantly higher than that of the low-arousal

group ($M = 3.15$, $SD = 1.13$), and the difference was statistically significant ($t(78) = 9.41$, $p < 0.001$). Furthermore, the test results showed no significant difference in the baseline understanding of AI instruction between the two groups of participants ($p = 0.918$), ensuring homogeneity between the experimental groups.

3.3. Measures

All core variables in this study were measured using a seven-point Likert scale (1 = "Strongly Disagree", 7 = "Strongly Agree"). The measurement items were derived from mature scales in existing literature and were appropriately adapted to the AI education context of this study to ensure the reliability and validity of the measurements.

3.3.1. Teaching Experience

This variable was comprehensively assessed through four items. Learning satisfaction was adapted from the user satisfaction scale by Wang et al.(2019a); perceived learning effectiveness and continuance intention referenced the evaluation of arousal usefulness by Yin et al. (2017) and the study on AI education acceptance by Xie et al. (2025); perceived learning effectiveness focused on assessing learners' subjective sense of gain and self-efficacy confirmation. Given the limited duration of the experimental stimulus materials, this study focused on measuring students' subjective certainty regarding knowledge reception rather than objective examination scores, to reflect learners' self-efficacy immediately following the interaction. Representative items included: "I feel that this teacher's explanation made it easy for me to understand these core concepts" and "I feel that I have effectively absorbed the core content of the video." In this study, the Cronbach's alpha coefficient for this scale was 0.73.

3.3.2. Psychological Distance

As the core mediating variable of this study, psychological distance was measured using three items. The scale was primarily based on the psychological distance theoretical framework of Trope and Liberman (2010) and specifically referenced adapted versions by Zheng et al.(2020) and Chen et al.(2024) in human-computer interaction and crisis response contexts. It focused on assessing the "social proximity" and "emotional understanding" perceived by students. Representative items included: "I feel that this AI teacher is very 'close' to me, like a kind partner" and "I feel that this AI teacher can understand my learning difficulties and needs." In this study, the Cronbach's alpha coefficient for this scale was 0.90.

3.3.3. Emotional Resonance

This variable was measured using three items. The scale originated from the narrative transportation scale by Escalas and Stern (2003) and was fine-tuned in combination with the research of Zhang and Bowman (2022) on rich storytelling promoting science learning.This dimension aimed to assess whether the AI teacher's language style stimulated students' emotional contagion and intrinsic interest. Representative items included: "I was infected by this AI teacher's explanation and felt an emotional resonance" and "I feel that this AI teacher's explanation style made me more interested in the learning content." In this study, the Cronbach's alpha coefficient for this scale was 0.90.

3.3.4. Concentration

This variable was measured using three items, primarily adapted from the learning concentration scale by Li and Yang (2016), while also referencing relevant indicators from Berger et al. (2023) regarding how linguistic features drive attention. This dimension assessed the degree of students' attention concentration and cognitive activity while watching the video. Representative items included: "During the lecture, my attention was very focused and did not wander" and "I actively thought about the questions raised and the content explained by the teacher." In this study, the Cronbach's alpha coefficient for this scale was 0.84.

3.3.5. Control Variables

To exclude potential interference from individual differences on the experimental results, based on suggestions from Huang and Rust (2018), participants' gender, age, education level, and prior understanding of AI instruction (measured by a single item) were included as control variables in the statistical analysis model.

Prior independent samples t-tests and Chi-square tests showed no significant differences between the high- and low-arousal groups on the aforementioned demographic variables ($p > 0.05$), demonstrating the validity of the random assignment in the experiment. Data analysis indicated a KMO value of 0.915, and Bartlett's test of sphericity was significant ($\chi^2 = 985.42$, $p < 0.001$), indicating suitability for factor analysis.

3.4. Data Analysis Procedure

All statistical analyses in this study were conducted using SPSS 27.0. First, the pre-experiment data underwent manipulation validity checks (independent samples t-test), participant homogeneity tests, and scale reliability tests. Upon confirming the qualification of the experimental materials, the formal experiment data were cleaned and tested for common method bias. The suitability of the analysis was simultaneously verified through KMO, Bartlett's test of sphericity, and Cronbach's α coefficients. Subsequently, descriptive statistics and Pearson correlation analysis were used to outline the relationships between variables. Independent samples t-tests were employed to analyze the inter-group differences between the high- and low-arousal groups, with Cohen's d effect sizes reported. Finally, the mediating role of psychological distance was verified using the Baron and Kenny stepwise regression method, controlling for variables such as gender and age. All analyses used $p < 0.05$ as the threshold for statistical significance.

4. Data Analysis and Results

4.1. Pre-experiment and Validity Check of Stimulus Materials

To ensure that the experimental manipulation of the "AI teacher's language arousal" successfully induced the intended psychological perceptions and to eliminate confounding factors prior to the formal experiment, a pre-experiment was conducted.

4.1.1. Pre-experiment Design and Participants

A total of 80 participants were recruited for the pre-experiment. Using a random assignment design, they were equally distributed into a high-arousal group ($n = 40$) and a low-arousal group ($n = 40$).

4.1.2. Manipulation Validity and Reliability Analysis

The study employed a four-item "Perceived Arousal Scale" to examine the validity of the stimulus materials. The data analysis results are presented in Table 1:

Manipulation Check: An independent samples t-test indicated that the comprehensive perceived score of the high-arousal group ($M = 5.39$, $SD = 1.00$) was significantly higher than that of the low-arousal group ($M = 3.15$, $SD = 1.13$), with the difference being highly statistically significant ($t(78) = 9.41$, $p < 0.001$). This demonstrates that the manipulation of language arousal levels in the experimental materials was highly successful and possessed good discriminability.

Homogeneity Test: There was no significant difference between the two groups regarding their "basic understanding of AI" ($t = 0.10$, $p = 0.918$). This confirms good homogeneity between the groups and rules out the potential interference of participants' prior knowledge.

Reliability Test: The overall Cronbach's alpha coefficient for the Perceived Arousal Scale was 0.908, indicating that the measurement tool possesses excellent internal consistency reliability.

Table 1. Descriptive Statistics and Difference Tests for Pre-experiment Items (N=80)

Measurement Dimensions and Items	High Arousal Group (n=40) M	Low Arousal Group (n=40) M	t	p
Perceived Arousal (Overall)	5.39	3.15	9.41	< 0.001
Marked intonation variation, rich in emotion	5.35	3.15	8.51	< 0.001
Rapid speech rate, full of passion	5.38	3.10	9.22	< 0.001
Calm, plain language style (Reverse-scored)	2.58	4.80	-8.87	< 0.001
Rational, highly logical expression (Reverse-scored)	2.58	4.85	-9.13	< 0.001
Control Variable: Understanding of AI	5.48	5.45	0.10	0.918

Note: Items 3 and 4 were reverse-scored.

4.2. Formal Experiment Sample and Data Quality

4.2.1. Sample Characteristics and Randomization Check

A total of 100 valid questionnaires were collected for this study. Adopting a randomized experimental design, participants were equally assigned to the high-arousal group (N=50) and the low-arousal group (N=50). To ensure the internal validity of the experimental results and exclude the confounding effects of individual participant characteristics on the experimental effects, this study first tested the homogeneity of baseline levels for demographic variables across both groups.

The statistical analysis results indicated the following:

Gender Distribution: The sample consisted of 47 males (47%) and 53 females (53%). Chi-square tests indicated no statistically significant difference in gender composition between the two groups (χ^2=0.64, p>0.05).

Educational Background: The participant population was predominantly undergraduate (85%), with 9% holding a master's degree or above, and 6% holding a junior college degree or below. Chi-square tests showed a balanced distribution of education between the two groups (χ^2=0.88,p>0.05).

Age and Experience: The age of the sample was concentrated in the 19–24 range. Independent samples t-tests further confirmed that there were no significant differences between the two groups in terms of mean age and prior understanding of AI instruction (p>0.05).

These results indicate that the baseline characteristics of the two groups were highly consistent prior to the experiment, satisfying the homogeneity assumption of random assignment. This provides a logical basis for attributing subsequent inter-group differences to the experimental manipulation variable.

4.2.2. Reliability and Validity Test of Measurement Tools

To ensure the robustness of the evidence, this study employed Cronbach's alpha coefficients to examine reliability, combining the Kaiser-Meyer-Olkin (KMO) measure with Bartlett's test of sphericity to examine construct validity.

Reliability Indicators: The analysis showed that the internal consistency coefficients for all latent variables were at extremely high levels. Specifically, the Cronbach's α coefficients for psychological distance (α=0.915), emotional resonance (α=0.887), and concentration (α=0.903) were all significantly higher than the psychometric standard of 0.70, indicating that the scales possess good measurement stability.

Validity Indicators: The suitability test for factor analysis showed that the KMO measure of sampling adequacy was 0.915, and Bartlett's test of sphericity reached a significant level (χ^2=985.42, p<0.001) .This indicates that there are good structural relationships among the variables, making them suitable for subsequent factor analysis and hypothesis testing.

4.3. Descriptive Statistics and Correlation Analysis of Variables

Before conducting causal inference, this study first examined the relationships among the main variables through a Pearson correlation coefficient matrix to preliminarily verify the logical associations between variables. The results are presented in Table 2.

Table 2. Descriptive Statistics and Correlation Coefficient Matrix of Variables

Variable	*M*	*SD*	1	2	3	4	5	6	7
Linguistic Arousal	0.50	0.50	1.00						
Psychological Distance	4.70	1.16	0.38**						
Emotional Resonance	4.28	1.22	0.44**	0.12	1.00				
Attentional Engagement	4.46	0.92	0.28**	0.07	0.08	1.00			
Perceived Learning Effectiveness	4.66	0.89	0.33**	0.39**	0.16	0.04	1.00		
Learning Satisfaction	4.33	1.07	0.31**	0.02	0.06	0.03	-0.01	1.00	
Continuance Intention	4.22	0.93	0.39**	0.09	0.16	0.18	-0.03	0.13	1.00

Note: N = 100, ** $p < 0.01$ (two-tailed test). Language arousal: 0 = low arousal, 1 = high arousal.

This study employed a Pearson correlation coefficient matrix to examine the relationships among the main variables (Table 2). The analysis results indicate that the independent variable, language arousal, is positively correlated with the dependent variable, perceived learning effectiveness ($r = 0.33$, $p < 0.01$). It also shows positive correlations with learning satisfaction ($r = 0.31$, $p < 0.01$) and continuance intention ($r = 0.39$, $p < 0.01$). Regarding the relationships between the independent variable and the mediating and process variables, language arousal showed significant positive associations with psychological distance ($r = 0.38$, $p < 0.01$), emotional resonance ($r = 0.44$, $p < 0.01$), and concentration ($r = 0.28$, $p < 0.01$). Furthermore, the mediating variable, psychological distance, was also positively correlated with perceived learning effectiveness ($r = 0.39$, $p < 0.01$). The statistical correlations among the aforementioned variables all reached significance, providing data support for the subsequent regression analysis and mediation effect testing.

4.4. Hypothesis Testing

4.4.1. Main Effect Test of Language Arousal on Instructional Outcomes (H1 and H3)

To test Hypothesis H1 (instructional outcomes) and H3 (process indicators), this study employed independent samples t-tests to analyze the significance of inter-group differences and introduced Cohen's d to quantify the substantive magnitude of the effects. The results are presented in Table 3.

Table 3. Difference Tests and Effect Size Analysis for High- and Low-Arousal Groups

Construct Dimensions	High Arousal Group (N=50) M	Low Arousal Group (N=50) M	t	p	*Cohen's d*	Interpretation of Effect Size
Learning Satisfaction	4.66	4.00	3.22	0.002	0.64	Medium Effect
Continuance Intention	4.58	3.86	4.19	<0.001	0.84	Large Effect
Perceived Learning Effectiveness	4.95	4.37	3.44	0.001	0.69	Medium Effect
Psychological Distance	5.14	4.26	4.07	<0.001	0.81	Large Effect
Emotional Resonance	4.82	3.74	4.90	<0.001	0.98	Large Effect
Attentional Engagement	4.72	4.20	2.93	0.004	0.59	Medium Effect

The analysis results indicate:

(1) Enhancement of Instructional Outcomes: The high-arousal group scored significantly higher than the low-arousal group on three core indicators: learning satisfaction, continuance intention, and perceived learning effectiveness ($p < 0.01$). Cohen's d values ranged between 0.64 and 0.84, indicating that the high-arousal language style exerts a medium-to-strong substantive impact on the improvement of instructional effectiveness. Thus, Hypothesis H1 is supported.

(2) Optimization of Process Experience: In terms of emotional resonance and concentration, the high-arousal group also demonstrated significant advantages ($t > 2.93, p < 0.01$), validating the

effectiveness of high-arousal language in eliciting positive emotions and maintaining cognitive attention. Thus, Hypothesis H3 is supported.

4.4.2. Mediation Effect Test of Psychological Distance (H2)

To further reveal the underlying mechanism by which language arousal affects instructional outcomes and to verify the mediating role of psychological distance, this study conducted a stepwise regression analysis following the classic procedure by Baron and Kenny (1986). The analysis results, with "perceived learning effectiveness" as the outcome variable after controlling for variables such as gender, age, and prior knowledge, are presented in Table 4.

Table 4. Regression Analysis of the Mediating Effect of Psychological Distance

Procedure	Model Specification	Independent Variable	β	t	p	R^2
Step 1	Y(Effectiveness) ~ X	Linguistic Arousal	0.328	3.44	0.001	0.108
Step 2	M(Distance) ~ X	Linguistic Arousal	0.380	4.07	<0.001	0.145
Step 3	Y(Effectiveness) ~ X+M	Linguistic Arousal	0.209	2.12	0.037	0.193
		Psychological Distance	0.315	3.20	0.002	

Results Analysis and Mechanism Explanation:

Total Effect Test (Model 1): Language arousal significantly and positively predicted perceived learning effectiveness ($\beta = 0.328$, $p = 0.001$), indicating that the overall impact of the independent variable on the dependent variable is significant, satisfying the prerequisite for mediation testing.

Path a Test (Model 2): Language arousal demonstrated a strong positive predictive effect on the mediating variable, psychological distance ($\beta = 0.380$, $p < 0.001$), suggesting that increased arousal levels effectively shorten psychological distance.

Establishment of Mediation Effect (Model 3): When psychological distance was included in the regression equation, its impact on perceived learning effectiveness remained significant ($\beta = 0.315$, $p = 0.002$). Meanwhile, the direct prediction coefficient of language arousal ($\beta = 0.209$) remained significant ($p = 0.037$), but the coefficient value decreased compared to Model 1 (dropping from 0.328 to 0.209).

The data results indicate that psychological distance plays a partial mediating role between language arousal and perceived learning effectiveness. This suggests that the high-arousal language style of AI teachers operates through a "dual-pathway" mechanism: on one hand, it

indirectly promotes learning effectiveness via the core psychological mechanism of "shortening psychological distance"; on the other hand, the emotional energy carried by language arousal itself produces a direct motivational effect. This mechanism highlights the complexity and necessity of emotional instructional communication in AI education. Although this study employed the classic stepwise regression method to intuitively demonstrate the path of coefficient changes, future research could combine Bootstrap methods (such as the PROCESS model) to construct confidence intervals for supplementary verification to further confirm the robustness of the results. Nevertheless, the significant changes in regression coefficients observed currently provide sufficient empirical support for Hypothesis H2.

4.5. Chapter Summary

Based on experimental data from 100 participants, this chapter systematically verified the research hypotheses using methods including descriptive statistics, t-tests, and regression analysis. Empirical results show that: (1) the data quality and the reliability and validity of the measurement tools meet psychometric requirements; (2) increasing the language arousal of AI teachers can significantly improve students' learning experiences and outcomes; and (3) psychological distance is the core mechanism explaining this enhancement effect. The research results provide robust empirical support for all hypotheses.

5. Discussion

5.1. Theoretical Implications

5.1.1. Deepening the Applicability of Language Expectancy Theory (LET) in AI Educational Contexts

Language Expectancy Theory posits that individuals hold established normative expectations regarding the language use of communicators in specific contexts (Burgoon et al., 2002). This study extends this theory from interpersonal communication to the field of human-computer interaction. Traditionally, machines are expected to be "rational and cold." However, this study found that when AI teachers exhibited a passionate, high-arousal language style, this "positive expectation violation" of the traditional machine image did not induce cognitive dissonance. Instead, students interpreted it as a signal of high-quality instructional investment (Wang et al., 2025). This finding corrects the concerns of some scholars that "high arousal might lead to cognitive overload," proving that in AI instruction, appropriate emotional energy infusion is a necessary condition for building "warm education," as such energetic expression precisely compensates for the lack of vitality inherent in a virtual identity (Li and Sung, 2021).

5.1.2. Establishing the Core Explanatory Power of "Psychological Distance" in Human-AI Instructional Interaction

While existing studies largely focus on social presence or immersion, this study innovatively introduces "Psychological Distance Theory" to explain AI instructional effectiveness. The findings reveal that high-arousal language, through its vividness and infectivity, transforms the originally abstract and distant digital human image into a concrete and approachable "learning

partner." This psychological "proximity" effectively lowers students' defense mechanisms, shifting knowledge transmission from "external indoctrination" to "internal acceptance" (Weidlich et al., 2024). Shortening psychological distance enhances perceived intimacy and trust, and this proximity directly determines perceived learning effectiveness, providing solid psychological evidence for the transition of AI from "tool-like intelligence" to "feeling intelligence" (Jones et al., 2017). Furthermore, the findings imply that language arousal may help mitigate the "Uncanny Valley Effect." When high-fidelity digital humans lack corresponding emotional expression, they can easily trigger user psychological discomfort (Wu et al., 2024); however, high-arousal language endows the virtual avatar with vitality matching its appearance, thereby correcting the cognitive dissonance of "human-like but not human" and promoting psychological acceptance.

5.1.3. Validating the Auxiliary Motivational Mechanism of Emotion and Attention

Although psychological distance plays a core mediating role, this study also confirmed significant improvements in emotional resonance and attention engagement in high-arousal instruction. This indicates that high-arousal language constructs a learning field optimized for both "cognition and emotion": on one hand, high-arousal language serves as a strong social cue, stimulating intrinsic motivation through emotional rendering (emotional resonance), with emotion playing the role of a "memory anchor" in learning (Barclay, 2020); on the other hand, it maintains cognitive alertness (attention) through variations in acoustic features (intonation modulation), making linguistic characteristics a key factor driving engagement (Berger et al., 2023). This combined mechanism ensures the occurrence of deep learning (Li and Yang, 2016).

5.2. Practical Implications

5.2.1. AI Teacher Design Should Shift from "Visual Realism" to "Vocal Emotionalization"

Current AI digital human development often excessively pursues the refinement of facial modeling while neglecting the "emotional adaptability" of vocal interaction (Tastemirova et al., 2022). This study suggests that educational technology developers should prioritize optimizing the emotional expressiveness of text-to-speech (TTS) synthesis over investing heavily in enhancing visual rendering realism. In the design of educational products, high-arousal voice models—characterized by modulated intonation, clear stress, and passion—should be configured as the default to avoid the psychological alienation caused by monotonous "robotic voices." Such a voice strategy can effectively improve users' perceived usefulness, thereby enhancing continuance intention (Xie et al., 2025).

5.2.2. Utilizing "Psychological Proximity Strategies" to Build New Teacher-Student Relationships

When utilizing AI to assist instruction, educators and course designers should consciously embed strategies to shorten psychological distance. The role of the AI teacher should shift from a "knowledge porter" to a "learning companion." Utilizing high-arousal language to build a foundation of trust—such as using first-person narratives and incorporating empathetic statements—can significantly shorten the "interpersonal distance" between students and virtual

teachers (Liu and Wang, 2024). Establishing this intimate relationship is key to improving student stickiness and completion rates in online education (Wang et al., 2019b).

5.2.3. Personalized Matching of Arousal Levels to Optimize Cognitive Experience

Although this study confirms the overall advantage of high arousal, in practical applications, systems should possess dynamic adjustment capabilities. The impact of language arousal may follow a non-linear relationship, where excessive arousal could be counterproductive (Yin et al., 2017). Future AI teachers should possess affective computing capabilities to dynamically adjust language styles based on the difficulty of the instructional content and real-time student feedback: increasing arousal to maintain attention when explaining dry theories, while appropriately adjusting it during segments requiring deep reflection, thus achieving "personalized" emotional instruction.

6. Conclusion

6.1. Research Summary

Positioned at the frontier of the intersection between artificial intelligence and educational psychology, this study systematically examined the mechanism of the influence of AI digital human teachers' language arousal levels (high vs. low) on college students' learning experiences and outcomes through a rigorous empirical experimental design. The results clearly indicate that the language strategy of AI teachers is not merely a formal embellishment but a critical variable determining the quality of online instruction.

Specifically, this study draws the following core conclusions:

(1) High-arousal language possesses significant instructional advantages. Compared to the traditional plain and calm style, AI teachers employing high-arousal language—characterized by modulated intonation and passion—demonstrated overwhelming superiority in enhancing students' learning satisfaction, continuance intention, and perceived learning effectiveness. This indicates that in human-AI instructional interactions, a "warm" voice is more effective than "rational" narration in stimulating learners' intrinsic motivation.

(2) Psychological distance is the core transmission mechanism. This study confirmed that language arousal does not act directly on cognitive outcomes but operates through the mediating variable of "psychological distance." High-arousal language successfully breaks the "cold barrier" between humans and machines. By establishing a sense of intimacy akin to quasi-social interaction, it lowers students' cognitive defenses, thereby promoting the deep internalization of knowledge. This finding explains, from a psychological perspective, why "anthropomorphic" voice can translate into actual instructional productivity.

(3) The synergistic effect of emotion and attention. The study also found that high-arousal language can effectively elicit emotional resonance and maintain high levels of attention engagement, providing new empirical support for the "integrated view of emotion and cognition" in multimedia learning.

6.2. Theoretical Contributions and Innovations

At the theoretical level, this study constructed and validated an integrated model of "language features—psychological mechanisms—instructional effects." Its primary innovative values are reflected in:

(1) Expanding the application boundaries of Language Expectancy Theory (LET). While previous studies mostly focused on the language styles of human teachers, this study extends the theory to the realm of AI interaction. It confirms that in digital educational contexts, users hold positive language expectations of "high arousal and high engagement" for AI as well, correcting the traditional stereotype that "machines should remain rational and neutral."

(2) Establishing the explanatory power of "psychological distance" in AI education. This study is the first to introduce psychological distance theory into the attribution analysis of AI instructional effectiveness, proving that shortening psychological distance is a stronger predictor of perceived learning effectiveness than merely enhancing technological visual realism. This provides a new theoretical lens for understanding the relationship between human-computer trust and knowledge transmission.

6.3. Practical Suggestions

Based on the above findings, this study proposes the following suggestions for the research, development, and application of AI educational products:

(1) Technical development should shift towards "affective computing." Future development of AI teachers should not remain stalled at creating the "shell" of visual images but should be dedicated to imbuing the "soul" through speech synthesis technologies. Developers are advised to establish voice generation models based on affective computing, enabling AI to output high-arousal speech capable of conveying emotion to enhance instructional appeal.

(2) Instructional design should emphasize "psychological connection." When deploying AI teaching assistants, educators should consciously design interaction scripts capable of shortening psychological distance (e.g., using first-person narratives and increasing empathetic expressions). This transforms the AI from a simple "knowledge retriever" into a "learning companion," thereby improving completion rates and stickiness in online learning.

6.4. Limitations and Future Directions

Despite validating the significant promoting effect of high-arousal language on AI instructional outcomes through a rigorous experimental design, these "strong effect" results must be viewed with caution. Given the specificity of the experimental context, this study has the following limitations that need to be addressed in future research.

(1) Limitations regarding "ecological validity" and the "Hawthorne Effect" in laboratory environments. The data results of this study showed extremely high effect sizes, which should be attributed to some extent to the highly controlled laboratory environment. In closed, quiet, and interference-free experimental conditions, participants' attention was forcibly focused on the AI teacher, which may have artificially amplified the instructional effect of language arousal. In real

classroom environments, noise, multitasking, and peer interference could significantly weaken this "high-arousal advantage" (Kwon et al., 2024). Additionally, the "Hawthorne Effect" (performing better because of being observed) may have led to positively biased self-reported data. Future research should attempt to conduct field experiments on real online learning platforms (MOOCs) to verify the robustness of this strong effect in natural ecological settings.

(2) The "Novelty Effect" of short-term interaction may lead to overestimated results. This study only examined a single, short-term (18 seconds) learning experience. The extremely high satisfaction and attention observed may partly stem from the "Novelty Effect" produced when students first encounter a "high-arousal, anthropomorphic" AI teacher. Over time, the emotional dividend brought by novelty may fade rapidly, potentially leading to aesthetic fatigue or cognitive overload (Tastemirova et al., 2022). Therefore, the conclusions of this study are applicable only to the "initial contact" phase. Future research urgently needs to conduct longitudinal tracking designs to examine whether high-arousal language can sustain such high psychological proximity effects over a semester-long course.

(3) Boundary conditions related to sample homogeneity and disciplinary specificity. The sample in this study consisted entirely of college students (aged 19–24) with relatively strong cognitive abilities and high acceptance of digital technologies, and the experimental material selected was relatively accessible "Management" knowledge. This high "Task-Technology Fit" may be one of the reasons for the highly consistent data results. For groups with lower cognitive load tolerance (such as elementary school students or the elderly), or when learning highly difficult abstract disciplines (such as quantum mechanics), high-arousal language might not only fail to shorten psychological distance but could also become a source of interference (Zhang et al., 2025). Future research needs to further examine the boundary conditions of language arousal strategies across different cognitive style groups and disciplinary types.

(4) The potential impact of Common Method Bias. In this study, psychological distance, emotional resonance, and learning experience were all measured using self-report scales from the same group of participants at the same time point. Although statistical tests indicated good data reliability and validity, the single data source may have artificially inflated the correlations between variables (i.e., common method bias). Participants in the high-arousal group might have tended to give high scores on all positive indicators due to the "Halo Effect" (Podsakoff et al., 2003). Future research suggests employing multimodal data triangulation, such as combining eye-tracking data to objectively measure attention or using electroencephalography (EEG) indicators to physiologically characterize emotional arousal, thereby obtaining a more objective chain of evidence. Furthermore, this study only measured "perceived learning effectiveness"; although previous studies indicate that subjective perception and objective performance are usually positively correlated, they are not identical. Future research should introduce standardized post-tests to more accurately assess the actual impact of language arousal on knowledge retention and transfer.

Author Contributions:

All authors have read and agreed to the published version of the manuscript.

Funding:

This research received no external funding.

Institutional Review Board Statement:

Not applicable.

Informed Consent Statement:

Not applicable.

Data Availability Statement:

The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflict of Interest:

The authors declare no conflict of interest.

References

Barclay, K. (2020). Family, memory and emotion in the museum. Emotion, Space and Society, 35, 100679.

Berger, J., & Milkman, K. L. (2012). What makes online content viral? Journal of Marketing Research, 49(2), 192–205.

Berger, J., Moe, W. W., & Schweidel, D. A. (2023). What holds attention? Linguistic drivers of engagement. Journal of Marketing, 87(5), 793–809.

Burgoon, M., Pauls, V., & Roberts, D. L. (2002). Language expectancy theory. In J. Dillard, & M. Pfau (Eds.), The persuasion handbook: Developments in theory and practice (pp. 117–136). SAGE Publications Inc.

Chen, Y., Liu, X., & Zhang, J. (2024). Bridging the gap: The role of psychological distance in human-AI interaction in educational settings. Computers & Education, 198, 104825.

Escalas, J. E., & Stern, B. B. (2003). Sympathy and empathy: Emotional responses to advertising dramas. Journal of Consumer Research, 29(4), 566–578.

Huang, M. H., & Rust, R. T. (2018). Artificial intelligence in service. Journal of Service Research, 21(2), 155–172.

Jones, C., Hine, D. W., & Marks, A. D. G. (2017). The future is now: Reducing psychological distance to increase public engagement with climate change. Risk Analysis, 37(2), 331–341.

Kwon, H., Choi, Y., Zhao, X., Hua, M., Wang, W., Garaj, V., & Lam, B. (2024). How immersive and interactive technologies affect the user experience and cultural exchange in the museum and gallery sector. The International Journal of the Inclusive Museum, 18(1), 135–160.

Li, L., & Yang, S. (2016). Focus or distraction? The impact of mobile learning on students' concentration and academic performance. Journal of Educational Computing Research, 54(3), 325–345.

Li, T., & Sung, Y. (2021). The effects of social presence and psychological distance on user trust in chatbots. International Journal of Human-Computer Interaction, 37(13), 1234–1245.

Li, X., & Sung, Y. (2021). Anthropomorphism brings us closer: The mediating role of psychological distance in User-AI assistant interactions. Computers in Human Behavior, 118, 106680.

Li, X., & Yang, X. (2016). Effects of learning styles and interest on concentration and achievement of students in mobile learning. Journal of Educational Computing Research, 54(7), 922–945.

Liberman, N., & Trope, Y. (2014). Traversing psychological distance. Trends in Cognitive Sciences, 18(7), 364–369.

Liu, J., & Wang, S. (2024). A study of the effect of viewing online health popular science information on users' willingness to change health behaviors—Based on the psychological distance perspective. Current Psychology, 43(38), 30135–30147.

Ludwig, S., De Ruyter, K., Friedman, M., Brüggen, E. C., Wetzels, M., & Pfann, G. (2013). More than words: The influence of affective content and linguistic style matches in online reviews on conversion rates. Journal of Marketing, 77(1), 87–103.

Massara, F., & Severino, F. (2013). Psychological distance in the heritage experience. Annals of Tourism Research, 42, 108–129.

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. Journal of Applied Psychology, 88(5), 879–903.

Tastemirova, A., Schneider, J., Kruse, L. C., Heinzle, S., & Brocke, J. V. (2022). Microexpressions in digital humans: Perceived affect, sincerity, and trustworthiness. Electronic Markets, 32(3), 1603–1620.

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. Psychological Review, 117(2), 440–463.

Wang, B., Han, Y., Kandampully, J., & Lu, X. (2025). How language arousal affects purchase intentions in online retailing? The role of virtual versus human influencers, language typicality, and trust. Journal of Retailing and Consumer Services, 82, 104106.

Wang, L., Ye, Z., & Zhu, S. (2023). Research on college students' emotional experience in online learning. International Journal of Continuing Engineering Education and Life-Long Learning, 33(6), 523–535.

Wang, W.-T., Ou, W.-M., & Chen, W.-Y. (2019a). The impact of inertia and user satisfaction on the continuance intentions to use mobile communication applications: A mobile service quality perspective. International Journal of Information Management, 44, 178–193.

Wang, W.-T., Wang, Y.-S., & Liu, E.-R. (2019b). The stickiness intention of group-buying websites: The integration of the quality-value-satisfaction framework and trust. Electronic Commerce Research and Applications, 15, 100–115.

Weidlich, J., Yau, J., & Kreijns, K. (2024). Social presence and psychological distance: A construal level account for online distance learning. Education and Information Technologies, 29(1), 401–423.

Wu, H., Chen, Z., Huang, Y., & Tu, H. (2024). Research on the uncanny valley effect in artificial intelligence news anchors. Multimedia Tools and Applications, 83(1), 62581–62606.

Xie, C., Xie, Y., Wang, Y., Zhou, P., Lu, L., Feng, Y., & Liang, C. (2025). Understanding older adults' continued-use intention of AI voice assistants. Universal Access in the Information Society, 24(2), 1687–1699.

Xie, H., Chu, H.-C., Hwang, G.-J., & Wang, C.-C. (2025). Trends and development in artificial intelligence in education: A systematic review. Educational Technology & Society, 28(1), 1–25.

Yin, D., Bond, S. D., & Zhang, H. (2017). Keep your cool or let it out: Nonlinear effects of expressed arousal on helpfulness of online reviews. Journal of Marketing Research, 54(3), 418–436.

Zhang, L., & Bowman, D. A. (2022). Exploring effect of level of storytelling richness on science learning in interactive and immersive virtual reality. ACM International Conference on Interactive Media Experiences (IMX '22), 19–32.

Zhang, Z., Van Lieshout, L. L. F., Colizoli, O., Li, H., Yang, T., Liu, C., Qin, S., & Bekkering, H. (2025). A cross-cultural comparison of intrinsic and extrinsic motivational drives for learning. Cognitive, Affective, & Behavioral Neuroscience, 25(1), 25–44.

Zheng, L., Miao, M., & Gan, Y. (2020). Perceived control buffers the effects of the COVID-19 pandemic on general health and life satisfaction: The mediating role of psychological distance. Applied Psychology: Health and Well-Being, 12(2), 471–492.