scientific reports

OPEN

Matching embodied conversational agent appearance to message emotion enhances persuasion in eHealth

Chengkun Tang[1], Li Wang[2], Duanwei Pan[3✉], Hui Zhang[4] & Ying Fang[5]

This study examines how matching Embodied Conversational Agents' (ECAs) appearance to message emotional tone enhances eHealth persuasion through the lens of Elaboration Likelihood Model (ELM) and Social Cues Theory. Using Event-Related Potential (ERP) measurements with 42 participants, we found professional ECAs delivering neutral messages elicited neural signatures of reduced cognitive conflict (smaller N400) and increased attention (larger LPP), demonstrating central route processing through credible source cues. Conversely, positive messages paired with casual appearances leveraged peripheral route persuasion via social rapport. Behavioral data confirmed these patterns, showing highest persuasion when professional appearance aligned with neutral tone - a congruence effect explained by both theories. Results provide actionable insights for designing persuasive ECAs in healthcare contexts by strategically combining visual and textual cues to optimize either credibility or approachability based on communication goals. The integration of neural and behavioral measures offers novel evidence for how multimodal cue matching operates in digital health interventions.

Non-communicable diseases (NCDs) have become the leading cause of global mortality, accounting for 76% of deaths worldwide[1]. This dramatic epidemiological shift has created an urgent need for innovative health communication solutions. Electronic health (eHealth) platforms have emerged as a promising approach to deliver scalable health interventions, but their effectiveness is often limited by low user engagement and high attrition rates[2].

Embodied conversational agents (ECAs) offer a potential solution by simulating face-to-face health communication. These digital interfaces can provide personalized health advice and support through natural interactions[3]. However, a critical challenge remains: how to design ECAs that are truly persuasive in promoting health behavior change. While previous research has separately examined the effects of ECA appearance and message tone, little is known about how these elements interact to influence user perceptions and decisions. This gap in knowledge significantly limits our ability to create optimally persuasive health communication agents.

ECA in eHealth

ECAs are autonomous software entities designed for health communication, utilizing embodied interfaces to simulate face-to-face interaction[4]. Their effectiveness stems from the integration of verbal and non-verbal social cues - including gestures, facial expressions, and vocal prosody - which enhance user engagement through naturalistic responses[5]. Florence, WHO's first virtual health worker, provided smoking cessation content to help people develop cessation plans[6]. Clinical applications demonstrate their versatility across: (1) chronic disease management (e.g., diabetes, hypertension), (2) mental health interventions (e.g., depression, anxiety), and (3) preventive care (e.g., smoking cessation, physical activity)[7,8]. Notably, 72% of users prefer ECAs for lifestyle-related guidance over complex medical consultations[9], suggesting their persuasive potential hinges on optimizing both informational content (central processing) and social presence (peripheral processing) - a

[1]College of Humanities, Donghua University, Shanghai, China. [2]School of Journalism, Communication University of China, Beijing, China. [3]The School of Basic Education, Shanghai Institute of Visual Arts, Shanghai, China. [4]Department of Physical Education, Donghua University, Shanghai, China. [5]College of Humanities and Social Sciences, University of Science and Technology of China, Hefei, Anhui, China. ✉email: 87194931@qq.com

duality reflected in health communication models[10]. This evidence justifies our focus on lifestyle domains where ECAs show highest adoption rates.

Impact of ECA appearance features

The appearance of ECAs—encompassing attire, demographics, and expressions—serves as a critical social cue in health communication. Professional attire (e.g., white coats) consistently enhances perceived authority, replicating the physician "white coat effect" in digital contexts[11,12]. Recent evidence suggests this visual professionalism reduces cognitive conflict during message processing[13]. According to the Elaboration Likelihood Model (ELM), persuasion occurs through two distinct pathways: (1) the central route involves careful evaluation of message content when users are motivated and able to process information deeply; (2) the peripheral route relies on heuristic cues (like appearance) when users lack capacity or motivation for detailed analysis[14]. Here, professional attire serves as a strong peripheral cue that primes acceptance of health messages. Conversely, casual attire improves approachability but may undermine message credibility[15] —reflecting the fundamental warmth-competence tradeoff in social cognition, where warmth (approachability) and competence (expertise) are often perceived as inversely related[16].

Demographic features show similar contextual effects. Users prefer young female agents for peer-like interactions[17], whereas authority-driven scenarios benefit from mature appearances[18]. Expressive elements further modulate these effects: static professional images paired with neutral text achieve high persuasion by facilitating central route processing[19], while dynamic ECAs require emotional congruence between appearance and tone to optimize peripheral route effects[20]. This aligns with cue consistency principles in multimodal communication, where aligned visual and verbal cues enhance processing fluency[13].

Effects of the emotional tonality of health information

According to ELM, neutral tones facilitate central route processing by enabling objective evaluation of factual content, while positive tones leverage peripheral route persuasion through affective arousal[21]. This dichotomy is evident in health communication: neutral texts excel for complex medical information by enhancing credibility[20], whereas positive tones boost motivation in behavioral interventions through emotional resonance[22].

Notably, text-based emotional cues trigger stronger social perceptions than previously assumed. Users instinctively associate positive tones with communicative warmth and neutral tones with expertise[13]—a pattern explained by ELM's peripheral cue processing, where emotional valence serves as a heuristic for agent personality. However, this effect is context-dependent: while positive expressions enhance perceived agent helpfulness in peer interactions, neutral tones prove more effective for authority-driven advice[15]. This tension underscores the need for strategic tone selection based on communication goals.

The effect of matching ECA appearance to the emotional tone of the message on persuasion

Persuasion can be described as changing the attitudes and/or behavior of others. In the context of eHealth, the persuasive power of an ECA refers to its ability to effectively communicate health information and influence users' perceived attitudes, intentions, and behaviors. Studies have identified the important role of persuasion in that by improving the perceived persuasiveness of a system[23]. It has also been shown that ECAs with a match between appearance and health topics (e.g., Chef and Cooking) result in higher ratings of persuasion and intention to use[18].

Persuasion in eHealth requires strategic alignment between ECA appearance and message tone—a phenomenon where ELM's dual routes interact with Social Cues Theory's congruence principle[13]. When professional ECAs deliver neutral health messages, users engage central route processing to evaluate factual content, while the authoritative appearance simultaneously validates message credibility through peripheral cues[14]. Conversely, casual ECAs with positive messages leverage peripheral route persuasion through social rapport[24], demonstrating how warmth-competence tradeoffs[16] dictate optimal cue combinations.

Importantly, mismatches trigger cognitive strain—as predicted by ELM's principle that conflicting central/peripheral cues impair persuasion[25]. Current applications already reflect these insights: clinical ECAs like WHO's Florence use professional-neutral pairing for credibility[26], while lifestyle coaches adopt casual-positive combinations for engagement[18].

ERP in AI agents

Event-related potentials (ERPs), which are electrophysiological signals associated with neural responses to events, provides critical insights into AI agent interactions by capturing implicit neural processes that self-reports cannot access[27]. In ECA research, the N400 component (negative deflection presenting a peak 400 milliseconds after the stimulus appeared) reflects semantic congruence—directly measuring ELM's central route processing when users evaluate message-appearance matches[28,29]. Similarly, the LPP component (Late Positive Potential, a late positive component with a peak amplitude of about 600ms) indexes motivational attention allocation[30,31], quantifying peripheral route engagement through emotional arousal.

These neural markers may resolve key theoretical debates: (1) N400 amplitudes confirm professional-neutral pairings reduce cognitive conflict[32], validating Social Cues Theory's congruence principle; (2) LPP enhancements to positive-casual pairings demonstrate peripheral route efficacy[28]. Crucially, ERP data reveal ELM-predicted interactions between routes—when peripheral cues (appearance) and central content (tone) align, they synergistically enhance persuasion[29]. This explains why multimodal consistency—not isolated cues—drives ECA effectiveness[13].

Research objectives

While prior research has established the independent effects of ECA appearance and message tone, how their congruence influences persuasion through integrated neurocognitive pathways remains unexplored at the intersection of ELM and Social Cues Theory. This study aims to: (1) identify optimal appearance-tone matches for eHealth ECAs, (2) uncover the neural mechanisms underlying their persuasive effects, and (3) provide evidence-based design guidelines. Specifically, we test four hypotheses that examine both explicit evaluations and implicit processing:

H1 Users have higher perceptions of matching ECAs with a combination of neutral messages and professional images compared to other combinations.

H2 Users perceive ECAs with neutral emotional messages and professional appearance as more persuasive than ECAs with other combinations.

H3 Users perceive less conflict and produce smaller N400s when matched with neutral emotional text and professional image.

H4 Users perceive higher similarity and produce greater LPP when matched with neutral emotional text and professional image.

Materials and methods
Participants

We recruited 42 students (23 females and 19 males, mean age 20.95 ± 2.118 years) from a University as participants. While the use of a homogeneous sample (e.g., university students) may limit the generalizability of our findings, it helps to reduce between-subject variability and increase statistical power for detecting neurocognitive effects, which is particularly important for ERP research with its typically moderate effect sizes[33]. This approach allows for a more sensitive test of our experimental manipulation under controlled conditions, generalizability to broader populations requires future verification. All participants were right-handed, had normal visual acuity or corrected vision, had no history of neurological or psychiatric disorders. In addition, all participants were required to be well rested and not taking stimulants or psychotropic drugs. The study was approved by the Science and Technology Ethics Committee of a University, and all subjects signed an informed consent form before the experiment and were paid a certain amount of money at the end of the experiment.

Stimuli

Based on previous research[34], and avoiding interference from other variables, the ECAs are two cartoon female appearances with the same face shape, the same simple smile, and the same hairstyle, aged around 20–30 years old. Their appearance differs only in that one wears a white coat with a stethoscope and the other wears casual clothes. They were designed and generated by art and design professionals and discussed by three scholars specializing in communication and medicine.

The original health text materials included 18 text messages about healthy lifestyles (Fitness & Exercise, Healthy Eating, and Stress Management): a total of 9 positive texts and 9 negative texts were included, all written in Chinese. The three sub-themes were selected based on their prominence as modifiable risk factors in global health guidelines[35].The Textual materials were adapted from WebMD and Mayo Clinic—consumer health platforms with established medical content accuracy[36], then refined through a two-stage process: (1) a physician and a communication researcher classified texts as positive/neutral using CDC's Clear Communication Index[37] for clinical validity, and (2) a linguist standardized emotional tone using LIWC lexicon[38] to ensure linguistic consistency. The mood of the texts was manipulated through the following: positive emotional texts usually use optimistic and inspirational language, aiming to stimulate and elevate the reader's mood. Neutral tonal text, on the other hand, focuses on objective and factual expressions, avoids emotional overtones, and focuses on direct and precise delivery of information. Negative tones were excluded per WHO guidelines discouraging fear appeals in health promotion[39]. Therefore, information containing negative emotions is excluded from our study.

To confirm the validity of the stimulus material, we recruited online 31 university students (target users of eHealth interventions[40] who did not participate in the subsequent ERP experiment (15 females and 16 males, mean age 21 ± 1.4 years).first, rating ECA professionalism on a 5-point scale (1 = Very unprofessional to 5 = Very professional) based on explicit visual cues (white coat = professional, casual = unprofessional), then evaluating text emotionality using the same scale. Although non-experts, for this study, the perception of the target users (university students) is the most important criterion in and of itself. Paired t-tests demonstrated significant discriminability: professional vs. casual images ($p < 0.001$, d = 1.2); positive vs. neutral texts ($p < 0.001$, d = 1.1). The 6 most discriminative messages (3 positive/3 neutral) were selected as final stimuli. The formal experimental materials are shown in Fig. 1.

Procedure

This experiment was conducted in the laboratory of the College of Humanities, a University. The experimental apparatus was the MindBridge-NaNo (developed by Guangzhou Qianga Neuroscience Technology Co., Ltd.). According to the extended version of the International 10–20 Electrode Placement System (Fig. 2), the electrodes were located at 32 standard positions. Stimuli were presented on a 19-inch LCD (1920×1080 pixels, 60 Hz) screen that was sized for clear observation by participants (1920×1080 pixels).

Before the start of the experiment, all participants were seated 70 cm from the front of the computer screen to view the stimulus images, with a viewing angle of approximately 33° x 19°. The ERP task was programmed and presented using Python 3.8 software, and pictures containing textual information or appearances were repeated 10 times each, presenting the stimuli randomly to eliminate the sequential effect, as shown in Fig. 3. Initially, a 1-minute countdown was used to put subjects in a relaxed state; a "+" sign appeared to help them focus on the center of the picture, and then the formal experiment began. First, a random picture containing textual

Fig. 1. Experimental stimulus materials.

information appeared for 3000-5000ms, followed by an appearance of a picture containing an image of an ECA, and subjects were required to judge the persuasiveness of this ECA image about the textual information that had just appeared, with 1 indicating very unpersuasive and 5 indicating very persuasive. The stimuli were alternated like this, with a blank page with a "+" sign in the center before each stimulus appeared, for 400-600ms to bring the participants' visual perception back to baseline level. After the ERP experiment, subjects were required to fill out a questionnaire on the perceived persuasiveness of the stimuli and the degree of match. Each experiment lasted 30 min with 1 break in between.

Measurement

Participants' immediate responses to the ECA persuasion were investigated via keystrokes in the formal ERP experiment. In the post-test questionnaire, the stimulus material from the ERP experiment was presented to the participants again. For matching, participants were asked to select one of the two ECAs that they perceived as the best match through text messages of different emotions. For perceived persuasion, participants were asked to rate the following three questions (adapted from a validated scale[18]: (1) The health advice provided by this character is persuasive; (2) The health advice provided by this character will influence me; (3) The health advice provided by this character will make me pay attention to my own (this aspect of) health behavior. Ratings were all on a five-point Likert scale ranging from strongly disagree to strongly agree.

Data recording and analysis

EEG activity was recorded with a neuroscanning cap, and EEG signals were acquired at a rate of 1000 Hz. Reference electrodes (A1 and A2) were placed in the bilateral mastoid process, and the impedance of each electrode was less than 5 kΩ. After the recording was completed, offline preprocessing was performed using the EEGLAB toolbox to obtain clean ERP data. The preprocessing of the data consisted of the following steps: (1) Re-
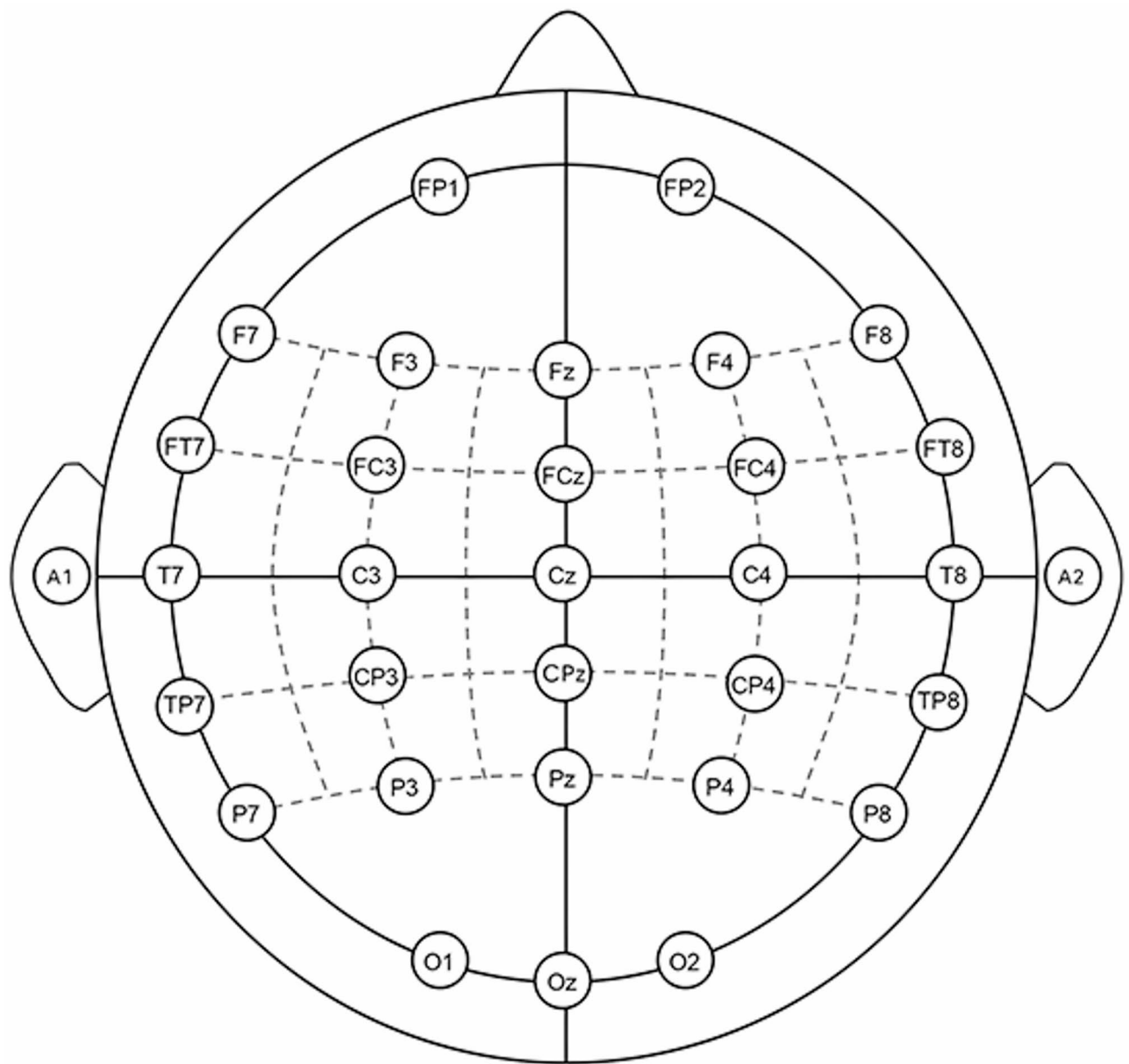
Fig. 2. A diagram of the electrodes used in the experiment.

referencing to the average of bilateral mastoids. (2) High-pass filtering at 30 Hz and low-pass filtering at 0.1 Hz; (3) Segmenting and baseline correction (-200ms to 800ms). (4) Independent Component Analysis (ICA); (5) Manual identification and artifact detection in epoched data; (6) Overlapping and averaging of ERP data.

Based on previous studies[29,41] and visual inspection of grand average waveform maps (Fig. 4), six electrodes in the frontal region and central frontal region (F3, FZ, F4, FC3, FCZ, FC4) and N400 (400-440ms) were selected for analysis, and nine electrodes in the region from the frontal lobe to the center (F3, FZ, F4, FC3, FCZ, FC4, C3, CZ, C4) were analyzed for LPP (620-670ms).

The keystroke persuasion data were analyzed using a two-way (appearance × emotional information) repeated ANOVA. For the ERP component, a three-way (appearance × emotional information × electrode) repeated ANOVA was used. Subjects' key press data and persuasion data from the post-test questionnaire were subjected to repeated ANOVA to determine whether subjects' perceived persuasion remained consistent across time points. Key press data and ERP amplitudes were compared using Pearson correlation analysis. All statistical analyses were tested for statistical significance using SPSS Statistics 26 and were considered statistically significant at $p < 0.05$. Data analyzed in SPSS were corrected using Greenhouse-Geisser.
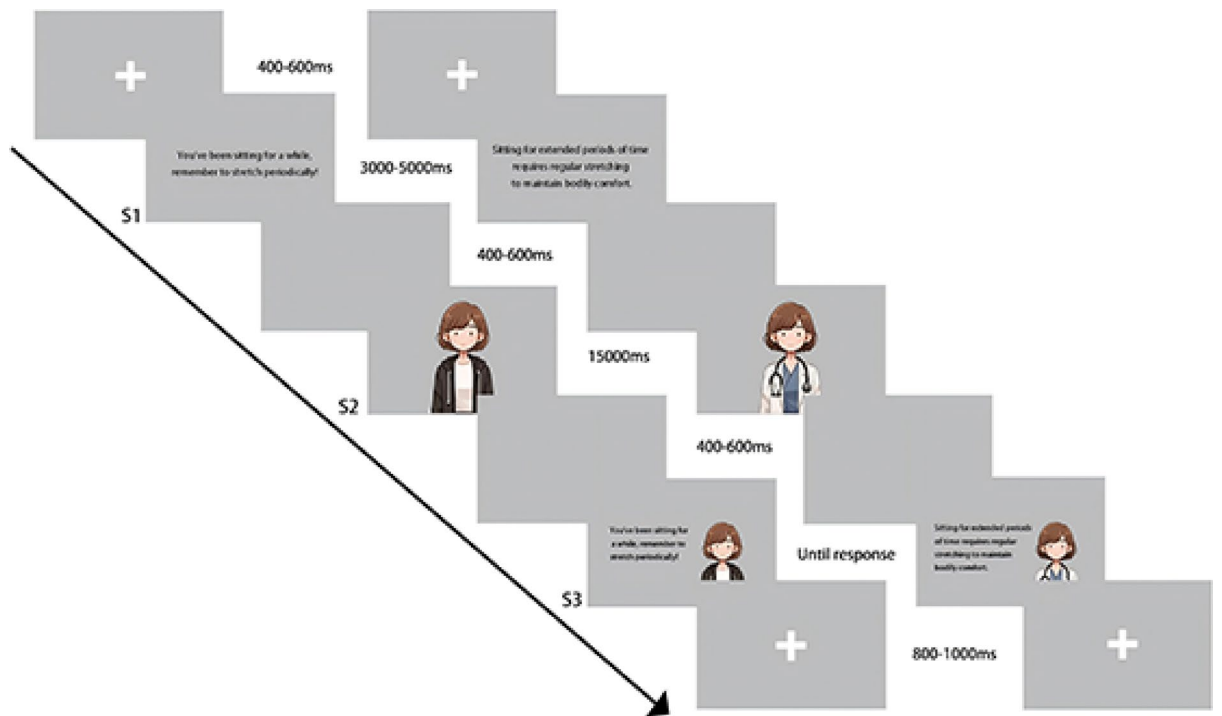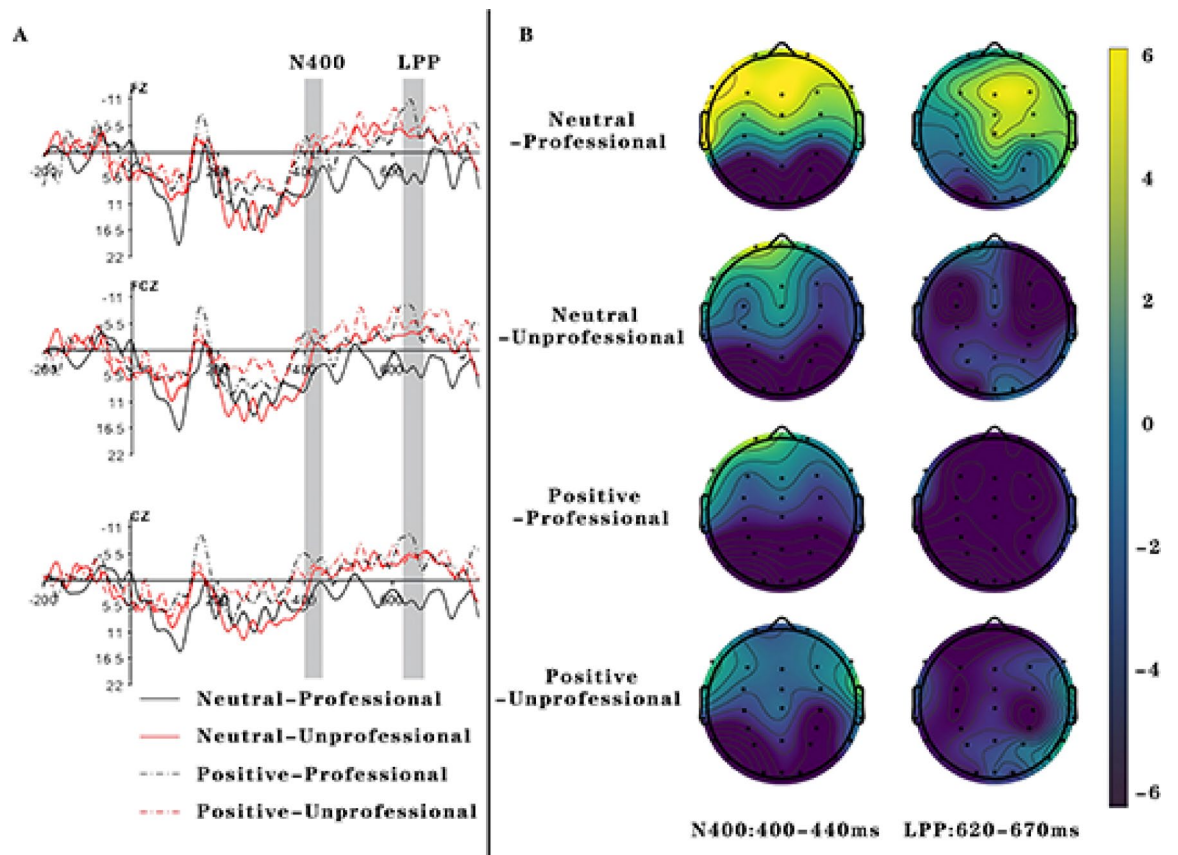
Fig. 3. Task paradigm with the timing of presentation.

Fig. 4. ERP waveforms and topographical maps of N400 and LPP under different conditions.

	Professional image	Unprofessional image	Total
Positive emotional text	39(92.9%)	3(7.1%)	42(100%)
Neutral emotional text	14(33.3%)	28(66.7%)	42(100%)

Table 1. Matched cognitive subject test under each eca's textual expression set.

	N400 (400–440 ms)			LPP (620–670 ms)			Persuasion		
	F	P	η2p	F	P	η2p	F	P	η2p
Image type	9.194	0.004	0.183	6.049	0.018	0.129	26.259	0.000	0.390
Emotion type	0.015	0.904	0.000	0.199	0.658	0.005	3.322	0.076	0.075
Image type*Emotion type	1.936	0.172	0.045	4.124	0.049	0.091	15.558	0.000	0.275

Table 2. Results of the repeated ANOVAs on the mean (amplitudes) of P2, LPP, and persuasion. Note: df = (1, 41); $p < 0.05$ is significant; $\eta^2 p$ is partial eta squared.

		N400 (400–440 ms)	LPP (620–670 ms)	Persuasion
Image type	Professional image	0.895(6.533)	− 1.986(5.952)	3.967(0.650)
	Unprofessional image	− 1.940(8.977)	− 4.231(8.195)	3.473(0.738)
Emotion type	Positive emotional text	− 0.594(8.295)	− 2.801(8.672)	3.639(0.728)
	Neutral emotional text	− 0.451(8.082)	− 3.416(7.051)	3.801(0.638)
Image type*Emotion type	Neutral *Professional	1.716(8.000)	− 1.324(7.858)	4.167(0.688)
	Neutral *Unprofessional	− 2.619(10.802)	− 5.507(9.260)	3.435(0.781)
	Positive *Professional	0.074(8.384)	− 2.649(8.391)	3.768(0.781)
	Positive *Unprofessional	− 1.261(10.139)	− 2.953(10.256)	3.511(0.839)

Table 3. Mean (amplitudes) of the ERP components and persuasion in the ERP study. Note: The number in the first row of every cell is the mean value. The standard error (S.E.) of each average is displayed in parentheses.

Results
Match
For neutral emotional text, 39 subjects (93%, Table 1) thought it was a better match to the professional image, and for positive emotional text, 28 subjects (67%) thought it was a better match to the professional image, so H1 was valid.

Difference between two persuasion judgments
In the ERP experiment, subjects' keystroke choices responded only to the question of whether they were persuasive or not, while the post-test questionnaire measured persuasion more specifically through the validated perceived persuasiveness scale. A repeated ANOVA was conducted to determine whether there were significant differences in subjects' persuasion judgments across time. The results showed a nonsignificant effect of time of measurement ($p = 0.57$), a nonsignificant interaction effect between time of measurement and other variables ($p > 0.05$), and a significant interaction effect only between emotional information and time of measurement ($p = 0.006$). However, on further analysis, the simple effect of emotional information was not significant either at the first measurement ($p = 0.08$) or at the second measurement ($p = 0.89$), and the simple effect of measurement time was not significant in either emotional information condition ($p = 0.08$; $p = 0.52$). That is, although the statistical tests showed a significant interaction effect, this effect may not be of much significance in practical applications. Thus, the results indicate that participants' evaluations of the persuasion of the stimuli remained consistent across time points.

ERP result
The results are shown in Tables 2 and 3. The effect of emotional information on N400 amplitude was not significant ($p = 0.90$), the main effect of image type was significant ($p = 0.004$), and the interaction effect of emotional information and image was not significant ($p = 0.17$). Unprofessional image (M = -1.940, SE = 8.977) elicited a larger N400 amplitude compared to professional image (M = 0.895, SE = 6.533). There was a nonsignificant effect of emotional information on LPP amplitude ($p = 0.66$), a significant effect of appearance type on LPP amplitude ($p = 0.02$), and a significant interaction effect ($p = 0.049$). Further analysis revealed that professional images had higher mean values of LPP amplitude than unprofessional images at the significance level ($p = 0.008$) in the neutral emotional text condition.

ERP results showed that neutral text combined with a professional image elicited smaller N400 and larger LPP amplitudes, supporting H3 and H4.

A two-way repeated ANOVA on push-button persuasion found (Tables 2 and 3) that there was no significant difference between the two emotional messages on persuasion ($p = 0.08$), a significant main effect of image type ($p < 0.001$), and a significant interaction between emotional message and appearance ($p < 0.001$). A simple effects analysis found that persuasion triggered by the professional image (M = 3.967, SE = 0.650) was significantly higher than persuasion triggered by unprofessional image (M = 3.473, SE = 0.738) ($p < 0.001$). A repeated ANOVA with the group as a single factor revealed significant differences in persuasion between combinations ($p < 0.001$). Bonferroni's multiple mean comparisons revealed that the combination of neutral emotional text and professional image was significantly more persuasive than the combination of neutral emotional text and unprofessional image ($p < 0.001$) and significantly more persuasive than the combination of positive emotional text and professional image ($p = 0.003$), and significantly higher than the positive emotional text and unprofessional image combination ($p < 0.001$). The other three groups were not significantly different from each other in terms of persuasion. Such results support H2.

Pearson correlation analyses of key press data and ERP amplitudes found no correlation between N400 and persuasion ($p = 0.23$), nor between LPP and persuasion ($p = 0.16$). Further subgroup analyses revealed that N400 showed a significant negative correlation with persuasion only in the combination of positive emotional text and professional image ($r = -0.306$, $p = 0.049$). In addition, in the positive mood condition, the elicited LPP and persuasion showed a significant negative correlation, both in combination with professional image ($r = -0.319$, $p = 0.04$) and unprofessional image ($r = -0.326$, $p = 0.04$).

Discussion
Principal results
This study examined how ECA appearances and emotional text affect perceived persuasion through the integrated lens of Social Cues Theory and the Elaboration Likelihood Model (ELM). The results demonstrated that professional ECAs matched with neutral text were consistently more persuasive (supporting H1 and H2), with neural evidence further validating these effects (H3 and H4).

The matching results revealed that 93% of participants perceived professional image ECAs as most neutral emotional health messages. This strong preference aligns with the "diagnosticity principle" in social cognition, where professional attire provides unambiguous expertise signals[26], while neutral tones maintain objective credibility. The white coat effect[11] explains how professional imagery automatically activates trust schemas, and when combined with fact-based messaging, creates a powerful persuasive synergy that can significantly influence health behaviors[18]. Interestingly, unprofessional images paired with positive emotional text also showed notable matching effects (67% approval), likely because casual appearances foster peer-like rapport[42] while positive tones enhance social engagement[20] - a combination particularly effective for motivational contexts.

Behavioral data clearly established the persuasive superiority of professional-neutral pairings (M = 4.17 vs. 3.47 for unprofessional-neutral, $p < 0.001$). This effect reflects ELM's central route processing, where credible sources facilitate deeper message elaboration. The interaction analysis revealed an important nuance: while professional images suffered persuasive penalties when paired with positive texts (due to role incongruence[15], casual appearances actually benefited from positive emotional tones through what term "peer affinity effects"[24]. This dichotomy perfectly illustrates Social Cues Theory's core premise - different social roles (expert vs. peer) demand distinct communication styles.

ERP findings provided neural validation of these effects. The professional-neutral combination elicited significantly smaller N400 amplitudes ($p = 0.004$), indicating reduced cognitive conflict during schema matching[43] - neural evidence for the white coat effect's automaticity. Concurrently, larger LPP amplitudes ($p = 0.02$) reflected enhanced motivational attention, suggesting this pairing successfully engages both automatic and controlled processing systems[44,45]. These neural signatures confirm that optimal persuasion occurs when visual and verbal cues mutually reinforce expected social scripts[26].

Notably, mismatched conditions revealed the costs of violating cue consistency. Professional ECAs delivering positive messages produced significant N400-LPP dissociation ($r = -0.32$, $p = 0.04$), reflecting the neural strain of reconciling conflicting expertise and warmth cues[46]. Behaviorally, these pairings showed 12% lower persuasion scores, demonstrating how cue incongruence undermines ELM's peripheral route effectiveness. Similarly, positive emotional texts unexpectedly generated negative LPP-persuasion correlations regardless of accompanying images, likely because health contexts trigger unique emotional processing patterns where excessive positivity may seem inappropriate[31].

The complete findings collectively demonstrate that ECA persuasion operates through dual pathways: professional-neutral pairings dominate through credibility-driven systematic processing (ELM's central route), while casual-positive combinations offer alternative peripheral route appeal via social-emotional engagement. This comprehensive account bridges Social Cues Theory's focus on multimodal congruence with ELM's processing depth continuum, providing both theoretical integration and practical design principles for health communication systems.

Strengths and limitations
The study's primary strength lies in its triangulation of behavioral and neural measures to decode ECA persuasion mechanisms—an approach that operationalizes ELM's dual-process theory through complementary explicit (questionnaire) and implicit (ERP) indicators. By capturing both conscious evaluations and subconscious processing, we bridge Social Cues Theory's focus on multimodal congruence with ELM's attention to processing depth. The replication of key effects across methods (e.g., professional-neutral superiority in both N400/LPP and persuasion ratings) enhances validity.

Three limitations warrant consideration: (1) Focusing solely on lifestyle topics (fitness, nutrition, stress) prioritizes ecological validity over breadth, though this aligns with ECAs' predominant use cases[10]. (2) While static ECAs control for animation artifacts, they preclude examination of dynamic emotion-expression matching—a key dimension in Social Cues Theory's extended framework[20]. (3) University students offer cognitive homogeneity for ERP but limit generalizability to age/education extremes, particularly relevant given ELM's emphasis on receiver characteristics[14].

These constraints suggest three theoretical-methodological synergies for future work: (1) dynamic ECAs to test ELM's peripheral cue flexibility, (2) cross-cultural samples to probe Social Cues' universality, and (3) clinical populations where central/peripheral route balances may differ.

Conclusions

This study advances our understanding of health communication by demonstrating how the strategic alignment of ECA appearance and message tone optimizes persuasion through complementary mechanisms. The robust superiority of professional ECAs delivering neutral messages confirms Social Cues Theory's diagnosticity principle while validating ELM's central route processing under high-credibility conditions. Conversely, the effectiveness of casual-positive pairings illustrates peripheral route persuasion through social-emotional engagement. Crucially, our neural evidence reveals that successful persuasion requires congruence between visual cues (which prime processing expectations) and textual tones (which fulfill those expectations)—a synthesis only possible through integrating both theoretical frameworks. These findings provide actionable guidelines: professional appearances should dominate fact-based health communication, while peer-like ECAs suit motivational contexts. Future research should explore how these principles generalize across cultures, health literacy levels, and dynamic interaction contexts.

Data availability

The datasets generated during and/or analysed during the current study are not publicly available due to the need to protect the privacy of the participants but are available from the corresponding author on reasonable request.

References

1. Organization, W. H. Noncommunicable diseases. https://www.who.int/health-topics/noncommunicable-diseases#tab=tab_1.
2. Nijland, N. Grounding eHealth: towards a holistic framework for sustainable eHealth technologies. (2011).
3. Bickmore, T. W. & Picard, R. W. Establishing and maintaining long-term human-computer relationships. *ACM Trans. Computer-Human Interact. (TOCHI)*. **12**, 293–327 (2005).
4. Ruttkay, Z., Dormann, C. & Noot, H. in Workshop on Embodied Conversational Agents held at the 2002 AAMAS Conference. 27–66Springer, (2004).
5. Guadagno, R. E., Blascovich, J., Bailenson, J. N. & McCall, C. Virtual humans and persuasion: the effects of agency and behavioral realism. *Media Psychol.* **10**, 1–22 (2007).
6. Loveys, K., Lloyd, E., Sagar, M. & Broadbent, E. Development of a virtual human for supporting tobacco cessation during the COVID-19 pandemic. *J. Med. Internet Res.* **25**, 9. https://doi.org/10.2196/42310 (2023).
7. Bickmore, T. in The Handbook on Socially Interactive Agents: 20 years of Research on Embodied Conversational Agents, Intelligent Virtual Agents, and Social Robotics Volume 2: Interactivity, Platforms, Application 403–436 (2022).
8. Provoost, S., Lau, H. M., Ruwaard, J. & Riper, H. Embodied conversational agents in clinical psychology: A scoping review. *J. Med. Internet Res.* **19**, 17. https://doi.org/10.2196/jmir.6553 (2017).
9. Wutz, M., Hermes, M., Winter, V. & Köberlein-Neu, J. Factors influencing the Acceptability, Acceptance, and adoption of conversational agents in health care: integrative review. *J. Med. Internet Res.* **25**, 28. https://doi.org/10.2196/46548 (2023).
10. Kramer, L. L., ter Stal, S., Mulder, B. C., de Vet, E. & van Velsen, L. Developing embodied conversational agents for coaching people in a healthy lifestyle: scoping review. *J. Med. Internet Res.* **22** (11). https://doi.org/10.2196/14058 (2020).
11. Brase, G. L. & Richmond, J. The white-coat effect: physician attire and perceived authority, friendliness, and attractiveness. *J. Appl. Soc. Psychol.* **34**, 2469–2481. https://doi.org/10.1111/j.1559-1816.2004.tb01987.x (2004).
12. Parmar, D., Olafsson, S., Utami, D. & Bickmore, T. & Acm. in 18th ACM International Conference on Intelligent Virtual Agents (IVA). 301–306Assoc Computing Machinery, (2018).
13. Rheu, M., Shin, J. Y., Peng, W. & Huh-Yoo, J. Systematic review: Trust-building factors and implications for conversational agent design. *Int. J. Hum. Comput. Interact.* **37**, 81–96 (2021).
14. Petty, R. E. & Cacioppo, J. T. *In Communication and Persuasion: Central and Peripheral Routes To Attitude Change* 141–172 (Springer, 1986).
15. Liew, T. W. & Tan, S. M. Social cues and implications for designing expert and competent artificial agents: A systematic review. *Telematics Inform.* **65**, 101721 (2021).
16. Fiske, S. T., Cuddy, A. J. & Glick, P. Universal dimensions of social cognition: warmth and competence. *Trends Cogn. Sci.* **11**, 77–83 (2007).
17. ter Stal, S., Kramer, L. L., Tabak, M., Akker, H. & Hermens, H. op den Design Features of Embodied Conversational Agents in eHealth: a Literature Review. *Int. J. Hum. Comput. Stud.* **138**, 22. https://doi.org/10.1016/j.ijhcs.2020.102409 (2020).
18. Kramer, L. L., van Velsen, L., Mulder, B. C., ter Stal, S. & de Vet, E. Optimizing appreciation and persuasion of embodied conversational agents for health behavior change: A design experiment and focus group study. *Health Inf. J.* **29** (20). https://doi.org/10.1177/14604582231183390 (2023).
19. Baylor, A. L. Promoting motivation with virtual agents and avatars: role of visual presence and appearance. *Philos. Trans. R Soc. B-Biol Sci.* **364**, 3559–3565. https://doi.org/10.1098/rstb.2009.0148 (2009).
20. Ter Stal, S., Jongbloed, G. & Tabak, M. Embodied conversational agents in eHealth: how facial and textual expressions of positive and neutral emotions influence perceptions of mutual Understanding. *Interact. Comput.* **33**, 167–176. https://doi.org/10.1093/iwc/iwab019 (2021).
21. Petty, R. E., Cacioppo, J. T. & Schumann, D. Central and peripheral routes to advertising effectiveness: the moderating role of involvement. *J. Consum. Res.* **10**, 135–146 (1983).
22. Kelders, S. M., Kok, R. N., Ossebaard, H. C. & Van Gemert-Pijnen, J. Persuasive system design does matter: A systematic review of adherence to Web-Based interventions. *J. Med. Internet Res.* **14**, 17–40. https://doi.org/10.2196/jmir.2104 (2012).

23. Drozd, F., Lehto, T. & Oinas-Kukkonen, H. in Persuasive Technology. Design for Health and Safety: 7th International Conference, PERSUASIVE 2012, Linköping, Sweden, June 6–8, Proceedings 7. 157–168 (Springer). (2012).
24. Rosenberg-Kima, R. B., Baylor, A. L., Plant, E. A. & Doerr, C. E. Interface agents as social models for female students: the effects of agent visual presence and appearance on female students' attitudes and beliefs. *Comput. Hum. Behav.* **24**, 2741–2756. https://doi.org/10.1016/j.chb.2008.03.017 (2008).
25. Petty, R. E. & Cacioppo, J. T. in *Adv. Exp. Soc. Psychol.* **19** 123–205 (Elsevier, 1986).
26. Loveys, K., Sebaratnam, G., Sagar, M. & Broadbent, E. The effect of design features on relationship quality with embodied conversational agents: a systematic review. *Int. J. Social Robot.* **12**, 1293–1312 (2020).
27. Dietrich, D. E. et al. Differential effects of emotional content on event-related potentials in word recognition memory. *Neuropsychobiology* **43**, 96–101. https://doi.org/10.1159/000054874 (2001).
28. de Visser, E. J. et al. Learning from the slips of others: neural correlates of trust in automated agents. *Front. Hum. Neurosci.* **12**, 15. https://doi.org/10.3389/fnhum.2018.00309 (2018).
29. Wang, C. C., Li, Y. Y., Fu, W. Z. & Jin, J. Whether to trust chatbots: applying the event-related approach to understand consumers? Emotional experiences in interactions with chatbots in e-commerce. *J. Retail Consum. Serv.* **73**, 11. https://doi.org/10.1016/j.jretconser.2023.103325 (2023).
30. Chen, M. L. et al. The neural and psychological basis of herding in purchasing books online: an Event-Related potential study. *Cyberpsychology Behav. Soc. Netw.* **13**, 321–328. https://doi.org/10.1089/cyber.2009.0142 (2010).
31. Wang, Q., Meng, L., Liu, M., Wang, Q. & Ma, Q. How do social-based cues influence consumers' online purchase decisions? An event-related potential study. *Electron. Commer. Res.* **16**, 1–26 (2016).
32. Liu, J. H. & Mo, Z. The effects of review's mobile phone price on consumers' purchase intention: an Event-Related potential study. *J. Neurosci. Psychol. Econ.* **14**, 197–206. https://doi.org/10.1037/npe0000152 (2021).
33. Brysbaert, M. How many participants do we have to include in properly powered experiments? A tutorial of power analysis with reference tables. *J. Cognition.* **2**, 16 (2019).
34. ter Stal, S., Tabak, M., op den Akker, H., Beinema, T. & Hermens, H. Who do you prefer? The effect of age, gender and role on users' first impressions of embodied conversational agents in eHealth. *Int. J. Human–Computer Interact.* **36**, 881–892 (2020).
35. Organization, W. H. in Global action plan for the prevention and control of noncommunicable diseases 2013–2020 (2013).
36. Nowrouzi, B., Gohar, B., Nowrouzi-Kia, B., Garbaczewska, M. & Brewster, K. An examination of scope, completeness, credibility, and readability of health, medical, and nutritional information on the internet: A comparative study of Wikipedia, WebMD, and the Mayo clinic websites. *Can. J. Diabetes.* **39**, 71 (2015).
37. Control, C. f. D. & Prevention. (CDC).
38. Pennebaker, J. W., Boyd, R. L., Jordan, K. & Blackburn, K. The development and psychometric properties of LIWC2015. (2015).
39. Organization, W. H. in Guidelines on physical activity, sedentary behaviour and sleep for children under 5 years of age 36–36 (2019).
40. Baumel, A., Muench, F., Edan, S. & Kane, J. M. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. *J. Med. Internet Res.* **21**, e14567 (2019).
41. Liu, J. H. et al. The effect of reviewers' Self-Disclosure of personal review record on consumer purchase decisions: an erps investigation. *Front. Psychol.* **11**, 9. https://doi.org/10.3389/fpsyg.2020.609538 (2021).
42. Bailenson, J. N., Blascovich, J. & Guadagno, R. E. Self-Representations in immersive virtual environments. *J. Appl. Soc. Psychol.* **38**, 2673–2690. https://doi.org/10.1111/j.1559-1816.2008.00409.x (2008).
43. Zhang, W. K., Jin, J., Wang, A. L., Ma, Q. G. & Yu, H. H. Consumers' implicit motivation of purchasing luxury brands: an EEG Study. Psychol. *Res. Behav. Manag.* **12**, 913–929. https://doi.org/10.2147/prbm.S215751 (2019).
44. Schupp, H. T. et al. Affective picture processing: the late positive potential is modulated by motivational relevance. *Psychophysiology* **37**, 257–261. https://doi.org/10.1111/1469-8986.3720257 (2000).
45. Schupp, H. T., Markus, J., Weike, A. I. & Hamm, A. O. Emotional facilitation of sensory processing in the visual cortex. *Psychol. Sci.* **14**, 7–13 (2003).
46. Rasmussen, A. Electrophysiology of stereotypes: N400 as a measure of the beautiful is good stereotype. (2007).

Acknowledgements

Author contributions

Chengkun Tang and Duanwei Pan participated in the conception and design of the study. Chengkun Tang, Li Wang, and Duanwei Pan contributed to the methodology. Li Wang carried out data curation. Chengkun Tang, Li Wang, Ying Fang, and Hui Zhang conducted formal analysis and investigation, and participated in writing—review and editing of the manuscript. Chengkun Tang was responsible for writing—original draft preparation, visualization, supervision, and project administration. All authors read and approved the final manuscript.

Funding

Declarations

Competing interests

The authors declare no competing interests.

Ethical statement

The study was conducted in accordance with the Declaration of Helsinki, and approved by the Science and Technology Ethics Committee of Donghua University (protocol code: DHUEC-RW-2024-05).

Additional information

Correspondence and requests for materials should be addressed to D.P.

Reprints and permissions information is available at www.nature.com/reprints.