

Article

Not peer-reviewed version

---

# Human-like, Animal-like, or Object-like? The Impact of LLM-Based Virtual Doctor Avatar Design on User Emotion, Physiology, and Experience

---

[Han Zhang](#)<sup>\*</sup>, [Shiyi Wang](#), [Rui Peng](#)

Posted Date: 4 December 2025

doi: 10.20944/preprints202512.0332.v1

Keywords: virtual doctor avatars; large language models; user experience; emotional responses; social presence; digital mental health



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

# Human-like, Animal-like, or Object-like? The Impact of LLM-Based Virtual Doctor Avatar Design on User Emotion, Physiology, and Experience

Han Zhang \*, Shiyi Wang and Rui Peng

Intelligent Design, Platform at the Intersection of Art and Technology, Central China Normal University, Wuhan 430079, China

\* Correspondence: zhanghan120@ccnu.cn

## Abstract

Virtual agents powered by large language models are increasingly deployed in digital mental health services, yet the influence of avatar appearance on users' emotional, cognitive, and physiological responses remains insufficiently understood. This study examined how three avatar designs—animal-like, human-like, and object-like—shape affective experience, user evaluation, autonomic activity, and attentional allocation during virtual doctor interactions. Forty-two participants completed a within-subjects experiment involving self-reported affect ratings, multidimensional user-experience assessments, heart rate variability (HRV) measures, and eye-tracking indicators. Avatar type did not significantly affect changes in positive or negative affect. However, physiological data revealed clear divergences. The animal-like avatar elicited the strongest parasympathetic activation, reflected by significant increases in RMSSD and HF power, whereas the object-like avatar produced a sympathetic-dominant response. Across six user-experience dimensions, the animal-like avatar consistently received the highest evaluations. Eye-tracking results showed faster first fixation and longer face-directed fixation duration for the animal-like avatar, indicating stronger social attention. The human-like avatar demonstrated slightly delayed initial fixation, consistent with subtle yet nonsignificant uncanny-valley tendencies. These findings underscore the critical role of avatar visual design in shaping emotional safety, engagement, and social processing in virtual mental-health interactions.

**Keywords:** virtual doctor avatars; large language models; user experience; emotional responses; social presence; digital mental health

---

## 1. Introduction

The rapid evolution of large language models (LLMs), including GPT-4, Gemini, and Llama, has significantly advanced the capabilities of conversational artificial intelligence and accelerated its integration into mental-health services (Lawrence et al., 2024; Guo et al., 2024). Beyond the functional, task-oriented chatbots of earlier decades, contemporary LLM-based agents are designed to sustain naturalistic dialogue, provide emotional support, and facilitate long-term relational communication (Haensch, 2025; Sakhrani et al., 2021). These developments have heightened the relevance of virtual conversational agents in psychological interventions, where access barriers, stigma, and a global shortage of mental-health professionals continue to limit service availability (Mohammed, 2023). Consequently, LLM-based virtual doctors are increasingly deployed to provide low-threshold, empathetic, and personalized support, offering an alternative pathway to timely emotional assistance (Cho et al., 2023; Wang et al., 2024). Given the rapid clinical uptake of such systems, understanding how design features shape user responses has become a central scientific and practical concern (Haensch, 2025).

Among the many design variables that shape human–AI interaction, the visual embodiment of a virtual agent—including its appearance, degree of anthropomorphism, and expressive cues—plays a critical role in users’ emotional experiences, trust formation, and sense of interpersonal safety (Thaler et al., 2021; Lim et al., 2024). According to the Computers as Social Actors (CASA) framework, users naturally apply social norms and interpersonal expectations to conversational agents, even when they are aware that the agent is artificial (Riek, 2015). In therapeutic contexts, such socio-cognitive responses are particularly meaningful because visual cues can facilitate (Lim et al., 2024) or hinder the establishment of emotional safety and perceived therapeutic alliance (Bilalpur et al., 2024).

Different forms of avatar embodiment elicit distinct psychological reactions. Human-like avatars may enhance social presence, but their effectiveness depends on the naturalness of their facial features and expressions (Thaler et al., 2021). When human-like avatars approach a high degree of realism but fail to fully match natural human facial qualities, users may experience discomfort associated with the Uncanny Valley, a phenomenon shown to reduce trust, suppress social gaze, and impair rapport (De Borst & De Gelder, 2015; Cheetham, 2011; Fraser et al., 2025). By contrast, animal-like avatars are often perceived as warm, safe, and emotionally approachable (Lim et al., 2024). Research from evolutionary psychology and human–animal interaction shows that animal figures can attenuate defensive behavior, lower anxiety, and promote emotional openness (Artemiou et al., 2021), suggesting that animal-like virtual doctors may enhance users’ willingness to express personal concerns (Sestino & D’Angelo, 2023; Schneider & Bubeck, 2025). Conversely, object-like avatars lack salient social cues such as faces and eyes, which may limit emotional engagement and reduce users’ willingness to relate to the agent as a social partner (Mazerant et al., 2025). Despite these theoretical distinctions, empirical comparisons of human-like, animal-like, and object-like virtual doctors in mental-health contexts remain scarce.

Visual attention provides a behavioral window into these psychological processes (Cheng et al., 2023; Capozzi & Kingstone, 2024). Eye-tracking research demonstrates that users’ facial fixations reflect social interest, emotional involvement, and trust (Bhattacharya & Gwizdka, 2018; Tsang, 2018; Kohout et al., 2023). Reduced or delayed attention toward an agent’s face can signal discomfort, threat appraisal, or early cognitive avoidance (Lim et al., 2024). These processes are particularly salient when interacting with near-human avatars (Thaler et al., 2021). However, research has yet to establish how different avatar embodiments systematically shape gaze patterns in therapeutic human–AI interactions.

Physiological responses offer another critical dimension. Heart rate variability (HRV) is a well-established index of autonomic nervous system regulation and emotional arousal (McCraty & Shaffer, 2015; Gullett et al., 2023). Higher HRV, particularly elevated RMSSD and HF power, reflects parasympathetic activation associated with calmness and emotional safety, whereas lower HRV indicates heightened stress or vigilance (Shaffer & Ginsberg, 2017; Kim et al., 2018; Stephenson et al., 2021). Studies in HCI and social robotics have shown that avatar appearance can modulate physiological responses: designs perceived as threatening or uncanny increase sympathetic activation, whereas warm and friendly designs support relaxation and stress reduction (Llanes-Jurado et al., 2024). However, the physiological effects of different virtual-doctor embodiments in LLM-mediated mental-health consultations remain underexplored.

Synthesizing behavioral, attentional, and physiological evidence reveals three major gaps. First, few studies have directly compared human-like, animal-like, and object-like virtual doctors, despite theoretical predictions that these avatars elicit distinct perceptions of warmth, threat, social presence, and emotional safety. Second, existing research seldom integrates multimodal indicators such as self-reported affect, PANAS scores, user-experience ratings, HRV, and gaze behavior within a single experimental paradigm. Third, the mechanisms linking visual embodiment to emotional and physiological responses, particularly the explanatory role of social attention, remain insufficiently understood in LLM-based mental-health interactions.

To address these gaps, the present study examines how three avatar types of influence users’ emotional experience, physiological states, and visual attention during interactions with an LLM-

based virtual doctor. Specifically, we ask: RQ1, whether avatar design affects emotional experience and user evaluation; RQ2, whether avatar design evokes distinct physiological patterns, especially HRV indicators related to emotional safety; and RQ3, whether avatar design induces different visual-attention patterns and whether these patterns help explain emotional or physiological differences.

Drawing on prior theoretical frameworks, we propose the following hypotheses.

H1: Avatar design significantly shapes emotional experience and subjective evaluation, with animal-like avatars eliciting the most positive emotional responses and user ratings, followed by human-like avatars, and object-like avatars receiving the lowest evaluations.

H2: Avatar design modulates physiological responses, such that animal-like avatars elicit greater parasympathetic activation (higher RMSSD, SDNN, and HF) compared with human-like or object-like avatars, while object-like avatars elicit higher LF/HF ratios indicative of elevated stress.

H3: Avatar design affects visual-attention patterns: animal-like avatars will elicit more facial fixations, longer fixation durations, and shorter first-fixation latencies; human-like avatars will exhibit early avoidance characterized by longer first-fixation latency; and object-like avatars will receive the lowest overall gaze allocation to facial regions.

By integrating subjective, physiological, and attentional measures within a unified experimental framework, this study advances understanding of how visual embodiment shapes user experiences with virtual clinicians and provides evidence-based guidance for designing emotionally supportive and trustworthy digital mental-health systems.

## 2. Materials and Methods

### 2.1. Participants

Before data collection, a power analysis was conducted using G\*Power 3.1 to determine the minimum required sample size. Following recommendations from Cohen and related methodological literature (Kang, 2021), a medium effect size ( $f = 0.25$ ) was selected, resulting in a required minimum of 35 participants.

A total of 42 university students completed all experimental tasks. Participants were between 19 and 27 years old ( $M = 21.10$ ,  $SD = 1.86$ ), including 16 males ( $M = 23.4$ ,  $SD = 3.5$ ) and 26 females ( $M = 21.6$ ,  $SD = 2.8$ ). All participants had basic computer literacy and no prior experience interacting with virtual doctors or digital mental-health interventions. Recruitment was conducted through university advertisements, social media, and personal communication channels. Prior to participation, individuals completed a mental-health screening questionnaire to ensure that none had severe psychological conditions that could be adversely affected by the study.

Ethical approval for this study was obtained from the institutional ethics committee (Approval No. CCNU-IRB-202306002a). All participants provided written informed consent and were informed of the study objectives, procedures, potential risks (e.g., emotional discomfort triggered by discussing personal stress), and their right to withdraw at any time. Data were anonymized and handled in accordance with the Personal Information Protection Law (PIPL) of China.

### 2.2. Materials

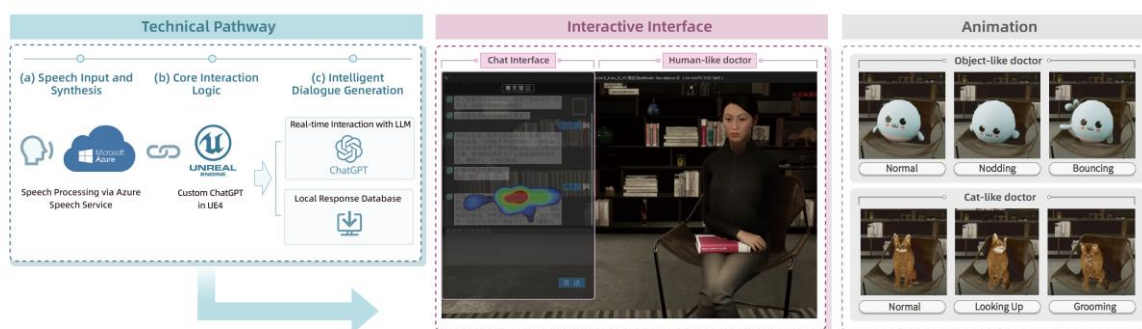
#### 2.2.1. Virtual Doctor Platform

The virtual doctor platform was developed using Unreal Engine 4 (UE4) and OpenAI's ChatGPT to enable natural language-based psychological consultation. As shown in Figure 1, the system integrates speech recognition, natural-language generation, and text-to-speech technologies to support seamless interaction through both voice and text (Cuadrado et al., 2020). Azure Speech Services handled speech-to-text processing, ChatGPT generated natural responses, and a text-to-speech module produced corresponding voice output. Mouth movements were synchronized with synthesized speech to enhance audiovisual coherence and user immersion (Bickmore & Picard, 2005).

Three avatar embodiments were created: a human-like doctor, a cat-like doctor, and an object-like doctor. Guided by the CASA framework, which suggests that people naturally attribute social characteristics to virtual agents (Nass et al., 1994; Xu et al., 2022), each avatar was designed to reflect a distinct social presence profile. The human-like avatar was modeled as an Asian female clinician to convey professionalism and warmth, consistent with literature suggesting that human-like avatars can enhance trust in therapeutic settings (Rehm et al., 2016; Jang et al., 2025). The animal-like avatar, inspired by a Chinese rural cat, was designed to reduce anxiety and foster a relaxed atmosphere, consistent with research on the comfort-enhancing role of non-human avatars (Ichikawa et al., 2023). The object-like avatar adopted a sphere-based design with eyes and hands to retain minimal social expressiveness while providing a non-threatening and playful experience.

Avatar modeling and animation were completed using MAYA, followed by animation blueprint configuration in Unreal Engine. Each avatar featured idle animations aligned with its persona. These behaviors included subtle book-reading animations for the human-like avatar, grooming movements for the cat avatar, and soft bouncing for the object-like avatar, enhancing social presence without introducing exaggerated or unnatural motion. (Oh et al., 2018; Sinatra et al., 2021).

The interaction environment was designed as a quiet, warmly lit consultation room furnished with books and comfortable seating to enhance trust and emotional comfort. Literature suggests that environmental cues significantly influence emotional safety and user trust in digital mental-health settings (Reuben et al., 2022; Lawrance et al., 2022; Werkmeister et al., 2024).



**Figure 1.** Virtual Doctor Platform.

### 2.2.2. Measures

**User Experience Scale:** User experience was assessed using a 24-item multidimensional User Experience Scale specifically designed to capture participants' perceptions of the virtual doctor during interaction. The scale comprises six dimensions, each containing four items: Emotional Resonance, Competence, Presence, Affiliation, Appearance Evaluation, and Satisfaction. All items were rated on a five-point Likert scale (1 = strongly disagree, 5 = strongly agree). The scale structure is consistent with validated frameworks commonly used in human-computer interaction and virtual agent research (Nowak & Biocca, 2003; Schuetzler et al., 2018). This scale demonstrated excellent internal consistency in the present study (Cronbach's  $\alpha = .93$ ), and each subscale showed acceptable reliability ( $\alpha$  values ranging from .82 to .90).

**Uncanny Valley Questionnaire:** To assess the potential uncanny effect of the human-like avatar, the Uncanny Valley Questionnaire (Mori et al., 2012) was administered only in the human-avatar condition. The 9-item scale evaluates perceived eeriness, human-likeness, and emotional distance, using a five-point rating format. Internal consistency for this scale was good (Cronbach's  $\alpha = .88$ ).

**Positive and negative Affect Scale (PANAS):** Participants' immediate emotional responses were measured using the Chinese version of the Positive and Negative Affect Schedule (Watson et al., 1988; Liu et al., 2020). The scale includes 20 items (10 positive, 10 negative), each rated on a five-point scale (1 = very slightly, 5 = extremely). Reliability was high in this study (Positive Affect  $\alpha = .89$ ; Negative Affect  $\alpha = .87$ ).

Eye movements were recorded using a Tobii Fusion Pro desktop eye tracker with a 120 Hz sampling rate, allowing continuous tracking of visual attention during the interaction with the virtual doctor. Based on the goals of social-cognitive analysis, two primary visual regions were defined as Areas of Interest (AOIs). Face AOI encompassed the avatar's facial region, including the eyes, mouth, and overall facial contour, which are essential for decoding social cues, interpreting emotional expressions, and assessing interpersonal intent as noted in previous social attention research (Itier & Batty, 2009). Text AOI included the dialog text bubbles and user-input area, capturing the extent to which participants allocated attention to conversational content during the interaction. For each AOI, three commonly used gaze indicators were extracted to characterize visual-attention patterns: fixation count, total fixation duration, and time to first fixation.

Heart rate variability (HRV) was collected using a Polar H10 chest-strap device at 250 Hz across the three-minute resting baseline (T1), the entire interaction period, and an immediate post-interaction measurement (T2). HRV analysis was performed in Kubios HRV Premium with automatic artifact correction using a medium threshold, following established psychophysiological guidelines (Tarvainen et al., 2014). RR intervals exceeding physiological plausibility or deviating substantially from adjacent intervals were flagged and corrected, and recordings requiring more than 10% correction were excluded to maintain signal reliability. Standard psychophysiological parameters were computed, including SDNN and RMSSD as time-domain indicators, LF and HF as frequency-domain indicators, and LF/HF as an index of autonomic balance. All preprocessing followed established guidelines for short-term HRV measurement (Laborde et al., 2017).

### 2.2.3. Experiments Procedures

The experiment followed a within-subjects design in which each participant completed interactions with all three avatars. To minimize order effects, avatar presentation was counterbalanced using a Latin-square arrangement. Upon arrival, participants signed consent forms, completed demographic materials, calibrated the eye tracker, and were fitted with the HRV sensor.

Before each avatar interaction, participants underwent a three-minute resting baseline (T1). Each interaction consisted of two phases. In the structured phase, participants discussed standardized psychological topics to ensure consistency in conversational content. In the free-interaction phase, participants engaged spontaneously with the avatar, and both eye-tracking and HRV data were continuously recorded. Immediately after each interaction, a short HRV measurement (T2) was taken without any intervening rest to capture the immediate autonomic response to the interaction. Participants then completed the PANAS, the User Experience Scale, and the Uncanny Valley Questionnaire when appropriate. A neutral low-arousal video lasting approximately five minutes was presented between conditions to minimize emotional carryover.

After completing all interactions, participants participated in a short open-ended interview to describe their preferences and perceptions. The full experiment lasted approximately ninety minutes per participant.

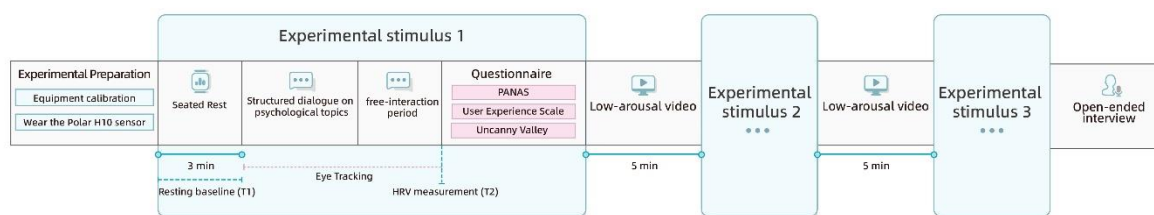


Figure 2. Experimental procedure.

### 2.2.4. Data Analysis

Data analyses were performed using SPSS 26.0 and Python 3.12.7. All variables were inspected for missing data, outliers, and distributional assumptions. Internal consistency of all questionnaire measures was assessed using Cronbach's  $\alpha$ .

Repeated-measures ANOVA was used to analyze the effects of avatar type on PANAS scores, user-experience ratings, uncanny valley responses, eye-tracking variables, and HRV indicators. Greenhouse–Geisser corrections were applied when sphericity assumptions were violated, and Bonferroni-adjusted comparisons were used to examine significant effects.

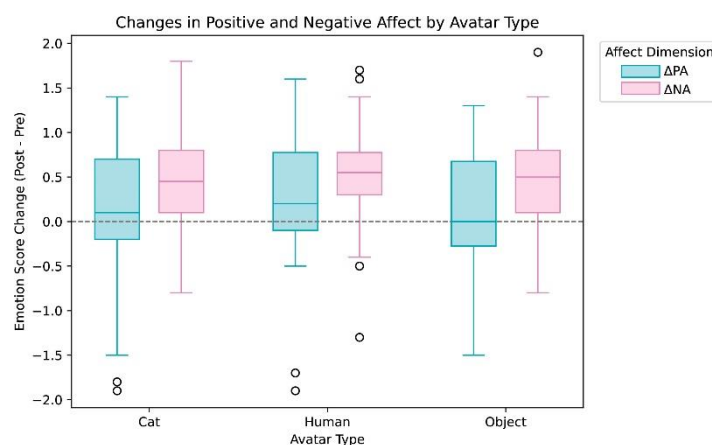
Eye-tracking analyses focused on contrasts among avatar types within each AOI, and HRV analyses examined both immediate post-interaction responses (T2) and their change from baseline (T1). Pearson correlations were computed to explore relationships among HRV, PANAS scores, and user-experience dimensions.

### 3. Results

#### 3.1. Effects of Avatar Type on Emotional Responses

##### 3.1.1. Subjective Emotional Responses

A repeated-measures ANOVA was conducted to examine whether avatar type (Cat, Human, Object) influenced participants' emotional responses. Changes in positive affect ( $\Delta$ PA) did not differ significantly across the three conditions,  $F(2, 82) = 2.06$ ,  $p = .135$ ,  $\eta^2p = .048$ . Likewise, changes in negative affect ( $\Delta$ NA) showed no significant effect of avatar type,  $F(2, 82) = 0.42$ ,  $p = .658$ ,  $\eta^2p = .010$ . Analyses of post-interaction affect scores yielded consistent patterns: neither post-interaction positive affect nor post-interaction negative affect varied significantly by avatar type (both  $F$ s = 2.06 and 0.42, respectively). As shown in Figure 3, distributions of  $\Delta$ PA and  $\Delta$ NA were largely overlapping across the Cat, Human, and Object avatars, indicating that the avatar's visual embodiment did not significantly modulate self-reported affect.



**Figure 3.** Boxplots of changes in positive affect ( $\Delta$ PA) and negative affect ( $\Delta$ NA) across the three avatar conditions (Cat, Human, Object).

##### 3.1.2. Physiological Responses

To assess whether avatar embodiment influenced autonomic activity, heart rate variability (HRV) indices were compared between baseline (T1) and the immediate post-interaction period (T2) across the three avatar conditions. Significant effects of avatar type emerged across multiple HRV metrics (Figure 4).

For SDNN, a general index of overall autonomic variability, a significant main effect of avatar type was observed,  $F(2, 82) = 11.92$ ,  $p < .001$ ,  $\eta^2 = .33$ . SDNN increased notably in both the Cat and Human conditions, with the Cat avatar producing the largest increase, whereas the Object avatar

elicited a decrease. Post-hoc comparisons indicated that SDNN was significantly higher in the Cat condition than in both the Human ( $p = .001$ ) and Object ( $p < .001$ ) conditions, suggesting a more relaxed physiological state during interactions with the Cat avatar.

RMSSD, a parasympathetic indicator, also revealed a significant effect of avatar type,  $F(2, 82) = 13.56$ ,  $p < .001$ ,  $\eta^2 = .36$ . RMSSD increased substantially in the Cat condition, whereas it showed only minimal change in the Human condition and a slight decrease in the Object condition. Post-hoc tests confirmed significantly higher RMSSD responses in the Cat condition relative to the Human ( $p = .001$ ) and Object ( $p < .001$ ) conditions.

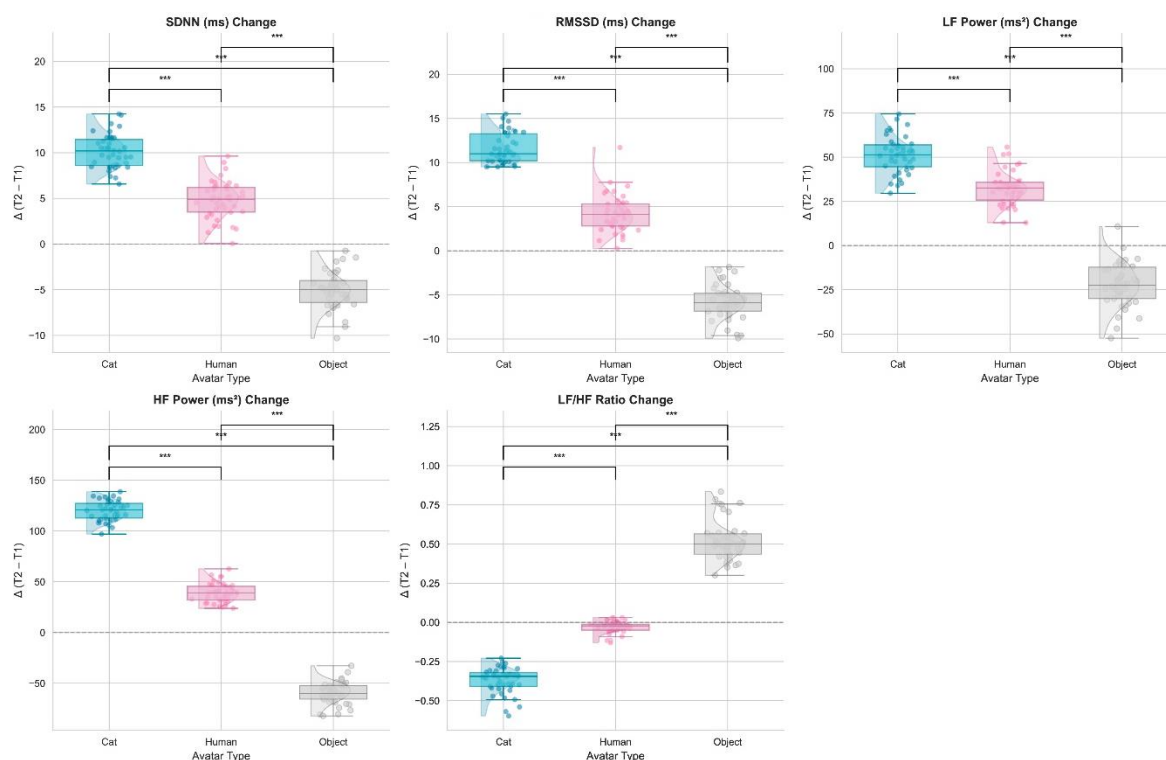
LF power, reflecting sympathetic activation, varied significantly across avatar types,  $F(2, 82) = 15.47$ ,  $p < .001$ ,  $\eta^2 = .38$ . LF gains were greatest under the Cat condition, significantly exceeding those in the Human ( $p = .001$ ) and Object ( $p < .001$ ) conditions. These increases suggest heightened—but non-stressful—arousal, consistent with adaptive engagement.

HF power, a parasympathetic index, also differed significantly among avatar types,  $F(2, 82) = 13.23$ ,  $p < .001$ ,  $\eta^2 = .35$ . HF increased most strongly in the Cat condition but declined slightly in the Object condition. Post-hoc comparisons indicated that HF was significantly higher in the Cat than in both the Human ( $p = .002$ ) and Object ( $p = .001$ ) conditions.

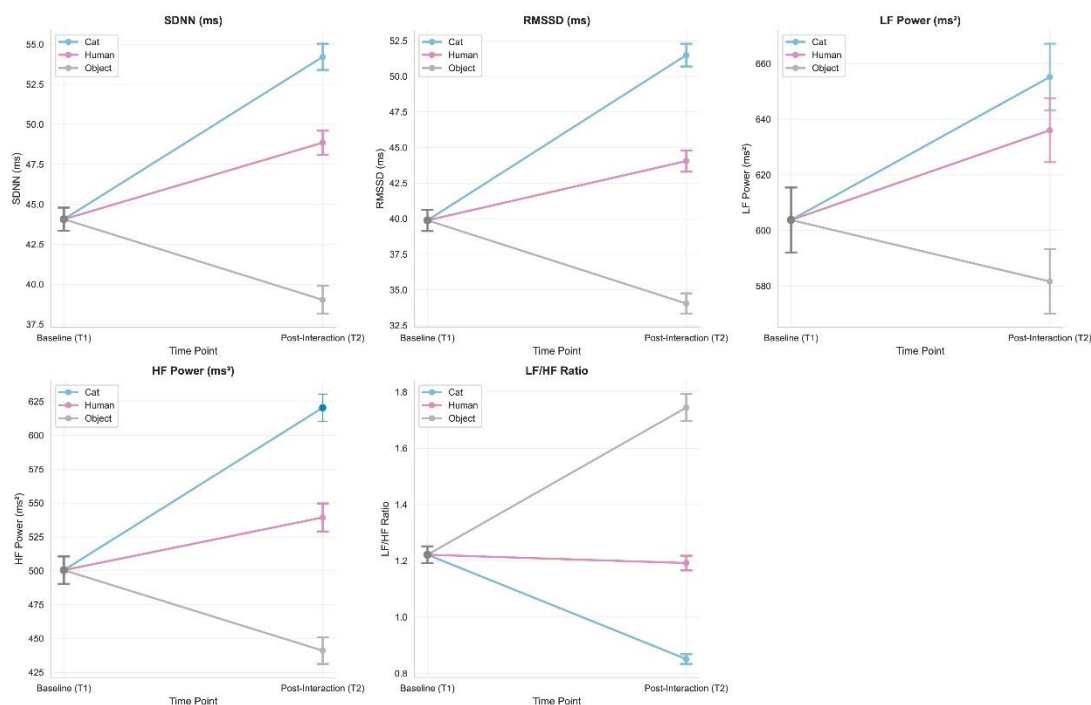
For LF/HF ratio, an index of sympathetic–parasympathetic balance, a significant effect of avatar type emerged,  $F(2, 82) = 10.02$ ,  $p < .001$ ,  $\eta^2 = .29$ . The Object condition produced the highest LF/HF ratio, indicating a sympathetic-dominant response and suggesting elevated stress. Significant differences were observed between Cat vs. Object ( $p = .001$ ) and Human vs. Object ( $p = .001$ ).

Further inspection of temporal changes (Figure 5) revealed distinct autonomic response patterns. The Cat avatar elicited robust increases in both parasympathetic (RMSSD, HF) and sympathetic (LF) activity from T1 to T2, reflecting a relaxed but engaged physiological profile. The Human avatar produced a mixed pattern, with increases in LF but modest reductions in RMSSD. In contrast, the Object avatar induced minimal HRV changes and an elevated LF/HF ratio, indicating comparatively higher physiological stress.

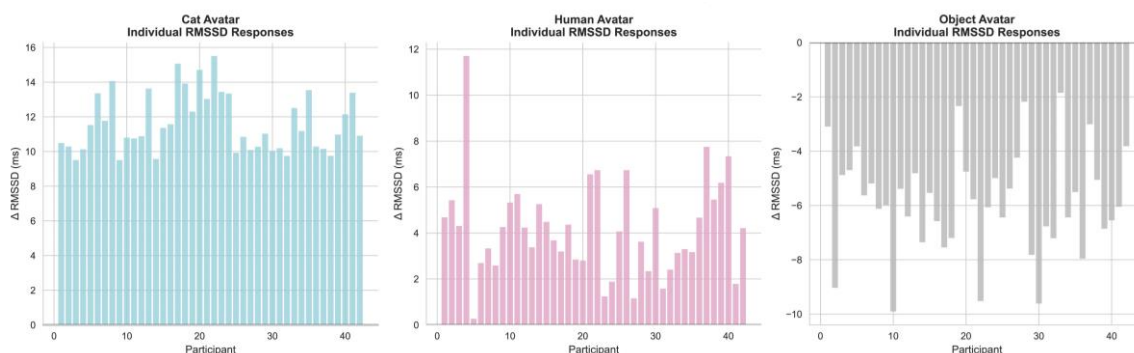
Individual response patterns (Figure 6) further highlighted these differences. Most participants displayed increased RMSSD in the Cat condition, whereas several participants exhibited pronounced RMSSD decreases in the Object condition, suggesting that the Object avatar may evoke stress responses in some individuals.



**Figure 4.** Box plots and scatter plots showing HRV changes (mean  $\pm$  SE) for SDNN, RMSSD, LF, HF, and LF/HF ratio from baseline (T1) to post-interaction (T2) across three avatar types (Cat, Human, Object). Repeated-measures ANOVA revealed significant differences between avatar conditions, with the Cat avatar eliciting the highest HRV responses. Asterisks indicate significant pairwise differences (\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ ).



**Figure 5.** Line plots showing HRV metrics (mean  $\pm$  SE) at baseline (T1) and post-interaction (T2) across three avatar types (Cat, Human, Object). Repeated-measures ANOVA revealed significant changes over time, with the Cat avatar showing the most pronounced HRV changes. Asterisks indicate significant pairwise differences (\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ ).



**Figure 6.** Bar plots showing individual  $\Delta$ RMSSD (T2 - T1) for each participant, with color-coded bars indicating increased (green) or decreased (red) parasympathetic activity. The Cat avatar condition primarily elicited parasympathetic activation, while the Object avatar prompted stress responses. Asterisks indicate significant differences between avatar conditions.

### 3.2. Effects of Avatar Type on User Experience

A series of repeated-measures ANOVAs was conducted to evaluate whether avatar type (Cat, Human, Object) influenced participants' evaluations across six user-experience dimensions: Emotional Resonance, Competence, Presence, Affiliation, Appearance Evaluation, and Satisfaction. As shown in Figure 7, consistent and robust effects of avatar type were observed across all

dimensions. Unless otherwise noted, all F statistics reflect uncorrected degrees of freedom; the application of Greenhouse–Geisser corrections did not alter the significance patterns.

For Emotional Resonance, avatar type exerted a significant main effect,  $F(2, 82) = 9.34$ ,  $p = .0002$ , partial  $\eta^2 = .19$ . Participants reported the highest emotional resonance with the Cat avatar ( $M = 3.38$ ,  $SE = 0.14$ ), followed by the Human avatar ( $M = 2.90$ ,  $SE = 0.12$ ), and the lowest ratings for the Object avatar ( $M = 2.68$ ,  $SE = 0.15$ ).

A similar pattern emerged for Competence,  $F(2, 82) = 6.58$ ,  $p = .0022$ , partial  $\eta^2 = .14$ , with the Cat avatar rated as most competent ( $M = 3.41$ ,  $SE = 0.13$ ), followed by the Human ( $M = 3.06$ ,  $SE = 0.11$ ) and Object ( $M = 2.84$ ,  $SE = 0.14$ ) avatars.

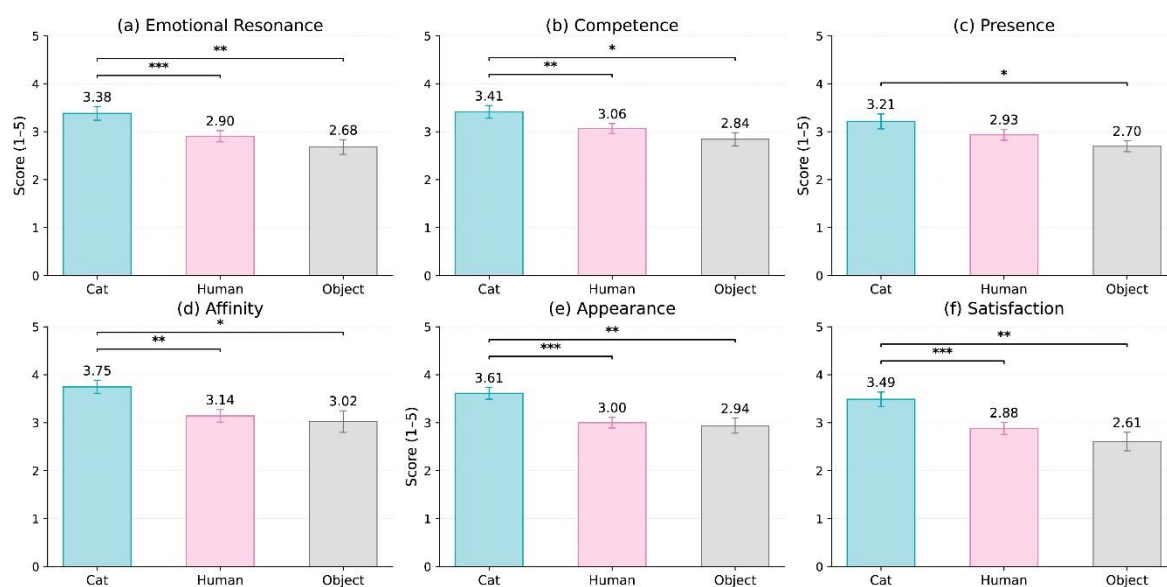
For Presence, the main effect of avatar type was significant,  $F(2, 82) = 5.32$ ,  $p = .0067$ , partial  $\eta^2 = .11$ . Presence ratings were highest for the Cat avatar ( $M = 3.21$ ,  $SE = 0.15$ ), moderate for the Human avatar ( $M = 2.93$ ,  $SE = 0.14$ ), and lowest for the Object avatar ( $M = 2.70$ ,  $SE = 0.15$ ).

The same directional trend was observed for Affiliation,  $F(2, 82) = 4.76$ ,  $p = .0110$ , partial  $\eta^2 = .10$ . Participants rated the Cat avatar highest in warmth and interpersonal affinity ( $M = 3.75$ ,  $SE = 0.14$ ), followed by the Human avatar ( $M = 3.14$ ,  $SE = 0.14$ ) and the Object avatar ( $M = 3.02$ ,  $SE = 0.23$ ).

For Appearance Evaluation, avatar type again yielded a significant effect,  $F(2, 82) = 8.47$ ,  $p = .00045$ , partial  $\eta^2 = .17$ . Participants rated the Cat avatar as the most visually appealing ( $M = 3.61$ ,  $SE = 0.12$ ), with substantially lower ratings for the Human ( $M = 3.00$ ,  $SE = 0.11$ ) and Object avatars ( $M = 2.94$ ,  $SE = 0.16$ ).

Finally, Satisfaction differed significantly across avatar types,  $F(2, 82) = 8.70$ ,  $p = .00037$ , partial  $\eta^2 = .18$ . Satisfaction was highest in the Cat condition ( $M = 3.49$ ,  $SE = 0.15$ ), followed by the Human ( $M = 2.88$ ,  $SE = 0.13$ ) and Object conditions ( $M = 2.61$ ,  $SE = 0.19$ ).

Across all six user-experience dimensions, a highly consistent pattern emerged: Cat > Human > Object, with moderate-to-large effect sizes (partial  $\eta^2$  approximately .10–.19). These results provide strong support for the hypothesis that animal-like avatars elicit more positive subjective experiences across emotional, social, and aesthetic domains compared to human-like and object-like avatars. Notably, the Human avatar did not show reliable advantages over the Object avatar, suggesting that moderate anthropomorphism does not necessarily guarantee superior user experience.



**Figure 7.** Comparison of user-experience ratings (mean  $\pm$  SE) across six evaluative dimensions for the Cat, Human, and Object avatars. Repeated-measures ANOVA revealed consistent effects of avatar type, with the Cat avatar rated highest on all dimensions. Asterisks denote significant pairwise differences ( $p < .05$ ,  $*p < .01$ ,  $p < .001$ ).

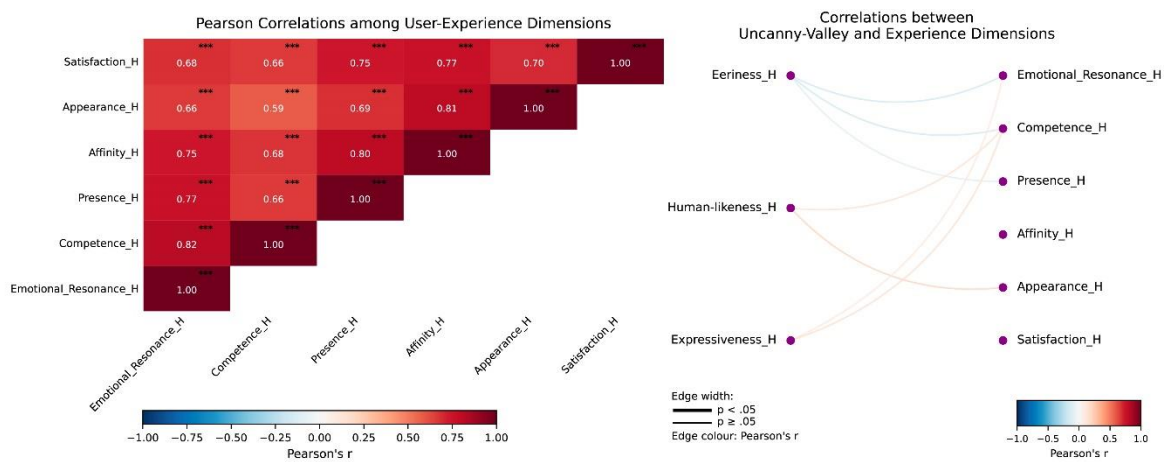
### 3.3. Uncanny Valley Effect in Human-like Virtual Doctors

To assess whether the human-like virtual doctor elicited an uncanny valley response, we examined the associations between three uncanny valley indicators—Eeriness, Human-likeness, and Expressiveness—and participants' user-experience ratings. Descriptive statistics indicated moderate levels of eeriness ( $M = 3.14$ ), relatively low human-likeness ( $M = 2.87$ ), and moderate expressiveness ( $M = 3.31$ ). These values suggest that the avatar was perceived as somewhat human-like but not sufficiently realistic to generate pronounced discomfort.

Correlation analyses revealed only weak and nonsignificant associations between uncanny valley responses and user-experience outcomes (Figure 8). Eeriness showed small negative correlations with Emotional Resonance ( $r = -.19$ ), Competence ( $r = -.19$ ), and Presence ( $r = -.11$ ), whereas Human-likeness and Expressiveness showed small positive correlations with Competence and Appearance ( $0.10 < r < 0.23$ ). None of these correlations reached statistical significance (all  $p > .23$ ), indicating that participants' subjective evaluations were largely independent of their uncanny valley perceptions.

Regression analyses further tested whether uncanny valley indicators predicted user experience. Across all models, standardized regression coefficients were small ( $|\beta| < .23$ ), and confidence intervals consistently crossed zero. Human-likeness showed a modest but nonsignificant positive trend in predicting Appearance ( $\beta = .228$ ,  $p = .146$ ), whereas Eeriness showed nonsignificant negative trends for Emotional Resonance ( $\beta = -.188$ ,  $p = .234$ ) and Competence ( $\beta = -.189$ ,  $p = .231$ ). These findings reinforce the conclusion that uncanny valley perceptions did not meaningfully influence evaluations of the human-like avatar.

Taken together, the results provide no strong evidence that the human-like virtual doctor elicited an uncanny valley effect. Although some effects were directionally consistent with uncanny valley theory—such as greater eeriness predicting lower social evaluations—the magnitudes were small, and none reached statistical significance. Overall, the human-like avatar appeared to fall short of the realism threshold necessary to trigger a pronounced uncanny valley response.



**Figure 8.** Correlation heatmap between uncanny valley indicators and user-experience dimensions for the human-like avatar.

### 3.4. The Impact of Visual Imagery on User Attention

To evaluate how avatar appearance influenced users' attentional allocation during interaction, we analyzed eye-tracking data across the three avatar types, namely Cat (animal-like), Human (human-like), and Object (object-like). Attention was examined within two Areas of Interest (AOIs), which were the facial region (Face AOI) and the dialog text region (Textbox AOI). The results demonstrated clear and systematic effects of avatar visual design on users' gaze behavior, as presented in Table 1 and Figures 9a-f.

For Face AOI, significant main effects of avatar type were found for all facial attention metrics. These included fixation count,  $F(2, 82) = 86.87$ ,  $p = 5.58 \times 10^{-21}$ ,  $\eta^2p = 0.609$ ; total fixation duration,  $F(2,$

82) = 249.23,  $p = 1.42 \times 10^{-35}$ ,  $\eta^2p = 0.789$ ; and time to first fixation (TFF),  $F(2, 82) = 235.01$ ,  $p = 1.11 \times 10^{-34}$ ,  $\eta^2p = 0.803$ . Post-hoc comparisons indicated that the Cat avatar received significantly more fixations and longer viewing durations than both the Human and Object avatars, with all  $p$  values below .001. The Object avatar consistently received the lowest level of facial attention.

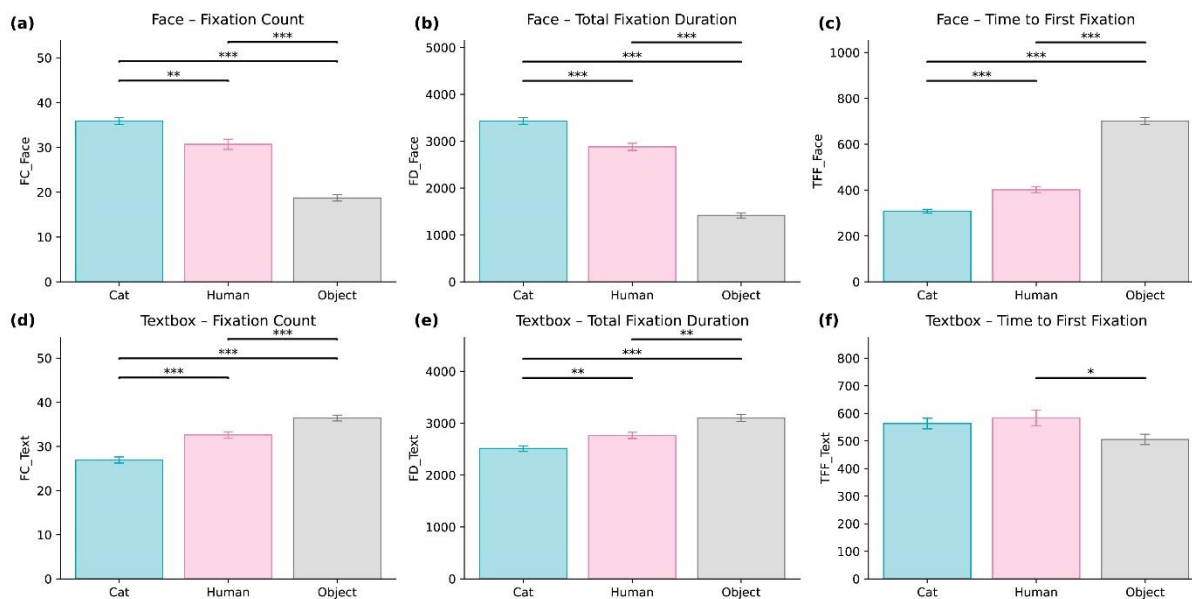
Although the Human avatar attracted more facial fixations than the Object avatar, it was associated with a longer TFF compared with the Cat avatar. This result suggests a delayed initial gaze response toward the Human avatar, a pattern that aligns with mild avoidance tendencies described in uncanny-valley research. Raincloud plots in Appendix Figure A1 further illustrate the strong separation among avatar conditions. The Cat avatar produced the most concentrated distribution of facial attention, while the Human and Object avatars produced wider variability. Kaplan–Meier survival curves in Appendix Figure A2 also showed that participants oriented to the Cat avatar’s face the fastest, followed by the Human avatar, with the Object avatar displaying the slowest orientation times.

**Table 1.** Bonferroni-corrected pairwise comparisons between avatar types for Face AOI and Text AOI measures.

Comparison	Face fixation count			Textbox fixation count		
	T(df)	p (Bonf.)	Hedges' g	T(df)	p (Bonf.)	Hedges' g
Cat vs Human	3.54 (41)	0.00304	0.815	5.17 (41)	3.34e-05	0.771
Cat vs Object	14.95 (41)	1.02e-17	3.442	9.15 (41)	5.32e-11	1.708
Human vs Object	8.74 (41)	1.99e-10	1.932	4.38 (41)		0.711
Comparison	Face fixation duration			Textbox fixation duration		
	T(df)	p (Bonf.)	Hedges' g	T(df)	p (Bonf.)	Hedges' g
Cat vs Human	6.74 (41)	1.16e-07	1.133	3.37 (41)	0.00414	0.617
Cat vs Object	19.85 (41)	3.69e-22	4.738	7.64 (41)	1.09e-08	1.436
Human vs Object	13.11 (41)	7.86e-16	2.412	4.27 (41)	2.92e-04	0.805
Comparison	Time to first fixation (face)			Time to first fixation (textbox)		
	T(df)	p (Bonf.)	Hedges' g	T(df)	p (Bonf.)	Hedges' g
Cat vs Human	-6.80 (41)	8.24e-08	-1.134	1.20 (41)	0.694	0.209
Cat vs Object	-22.18 (41)	1.02e-24	-4.857	-2.87 (41)	0.0166	-0.461
Human vs Object	-10.91 (41)	1.19e-12	-2.008	-1.76 (41)	0.264	-0.289

Avatar type also exerted a significant influence on attention directed toward the text region. Fixation count differed significantly across conditions,  $F(2, 82) = 55.15$ ,  $p = 6.67 \times 10^{-16}$ ,  $\eta^2p = 0.440$ , as did total fixation duration,  $F(2, 82) = 24.40$ ,  $p = 4.85 \times 10^{-9}$ ,  $\eta^2p = 0.286$ . Participants devoted more visual attention to the text region when interacting with the Cat avatar compared with both the Human and Object avatars, with all  $p$  values below .01.

For TFF in the Textbox AOI, avatar type showed a smaller but significant effect,  $F(2, 82) = 3.33$ ,  $p = .0408$ ,  $\eta^2p = 0.050$ . Post-hoc comparisons indicated that participants oriented to the text more rapidly in the Object condition than in the Cat condition ( $p = .016$ ). This suggests that when social cues are minimal, as in the Object avatar condition, users may shift their attention to textual information more quickly.



**Figure 9.** Eye-tracking metrics across avatar conditions for Face and Textbox AOIs. (a–f) Mean values ( $\pm$  s.e.m.) of fixation count, total fixation duration, and time to first fixation (TFF) for Cat, Human, and Object avatars across two areas of interest (AOI-Face and AOI-Textbox). Significant pairwise differences were assessed using Bonferroni-corrected paired t-tests, with asterisks indicating significance (\* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$ ). Error bars represent the standard error of the mean.

Across both AOIs, the Cat avatar consistently elicited the strongest and most rapid attentional engagement. The Human avatar produced moderate levels of facial engagement but showed delayed initial attention, consistent with subtle avoidance patterns. The Object avatar drew the least social attention and prompted the quickest transition toward text-based information, reflecting its low level of social cue richness.

These findings collectively indicate that visual design plays a critical role in shaping attentional strategies during digital mental-health interactions. Avatars with stronger socio-emotional cues, as illustrated by the Cat design, are more effective in capturing and sustaining users' visual attention.

#### 4. Discussion

The present study integrated subjective affect ratings, user-experience evaluations, physiological measures of autonomic activity, and eye-tracking indicators to investigate how different virtual doctor embodiments shape users' emotional, cognitive, and attentional responses during LLM-based mental-health interactions. Across these multimodal measures, two central observations emerged. First, avatar appearance exerted a systematic influence on social processing and attention allocation, although emotional self-reports and physiological reactions exhibited partially dissociated patterns. Second, the animal-like avatar consistently produced more favorable subjective and objective outcomes, whereas the human-like avatar elicited moderate social engagement without clear evidence of a strong uncanny valley response. This constellation of findings refines existing theoretical perspectives of human-agent interaction by suggesting that emotional safety and socio-affiliative affordances play a more decisive role than anthropomorphism alone in shaping responses to virtual clinicians.

A key contribution of this study concerns the divergence between subjective emotional reports and physiological responses. Although positive and negative affect ratings did not vary significantly across avatar types, HRV indices showed robust differences. Interactions with the animal-like avatar were associated with enhanced parasympathetic activation and more balanced autonomic regulation, whereas the object-like avatar was linked with reduced HRV and elevated sympathetic dominance. These results should be interpreted with caution, as HRV is influenced by multiple contextual factors

including cognitive load, posture, and speech-related respiration, which may partially contribute to the observed differences (Laborde et al., 2017). Even so, the dissociation between self-reported affect and HRV aligns with dual-systems perspectives suggesting that subjective feelings and physiological states can be governed by distinct regulatory mechanisms (Mauss & Robinson, 2009). The fact that participants were engaged in cognitively demanding linguistic interaction may have dampened conscious emotional awareness, whereas autonomic indices continued to reflect subtle variations in emotional safety or arousal associated with avatar appearance (Shaffer & Ginsberg, 2017). These findings underscore the value of incorporating psychophysiological measures such as HRV, electrodermal activity, or pupil dilation into future digital mental-health research, particularly when emotional shifts may be implicit or not accessible to self-report.

User-experience evaluations further demonstrated a robust and consistent pattern: the animal-like avatar was rated highest across emotional resonance, presence, appearance, and satisfaction, followed by the human-like avatar, with the object-like avatar receiving the lowest evaluations. While this pattern partially aligns with CASA predictions that more socially expressive agents elicit stronger social responses (Ho et al., 2018; Klein, 2025), the superior performance of the animal-like avatar suggests that anthropomorphism alone does not determine user experience. Instead, these results resonate with research on cross-species social affinity, which suggests that non-threatening animals can evoke strong feelings of safety, warmth, and social openness (Premathilake & Li, 2024). Such affective affordances may enhance emotional resonance and reduce social defensiveness—a desirable characteristic for mental-health applications. Consistent with findings from human–animal interaction literature (Beetz et al., 2012; Ahmed et al., 2024). The findings imply that cross-species social affinity and perceived emotional safety may constitute critical determinants of user comfort in virtual mental-health interactions, potentially overriding the effects of human resemblance.

By contrast, although the human-like avatar contained clear social cues, it did not provide experiential advantages over the animal-like design. Its moderate realism may not have been sufficient to activate human-like social schemas fully (S. Schneider et al., 2022; Dubois-Sage et al., 2023), while limitations in facial expressiveness or micro-behavioral dynamics may have introduced subtle inconsistencies. These features could result in mild perceptual discomfort, consistent with research showing that mid-realism avatars often lack the expressive coherence required to sustain trust and comfort (Baake et al., 2025; Huang et al., 2025). However, it is notable that the human-like avatar in this study did not elicit a strong uncanny valley effect. Correlation and regression analyses showed no significant relationships between eeriness, human-likeness, and user-experience ratings, suggesting that the avatar did not surpass the “realism threshold” required to evoke intense uncanny sensations (Mori et al., 2012; Thaler et al., 2021). Nevertheless, the delayed first fixation toward the human-like avatar’s face indicates subtle early perceptual hesitation, consistent with prior findings showing that early visual avoidance may occur prior to conscious reports of eeriness (Cheetham, 2011). Importantly, this avoidance effect was mild and did not translate into negative experiential evaluations.

Eye-tracking findings revealed that avatar appearance had a profound influence on attention allocation. The animal-like avatar attracted the highest number of fixations, the longest viewing durations, and the fastest initial attention toward the face, indicating strong social cue engagement. These results align with theories of social attention, which posit that faces with warm or affiliative features preferentially capture visual processing (End & Gamer, 2017; Kawakami et al., 2024). The human-like avatar produced moderate but less immediate social attention, reflecting the interplay between positive social cues and subtle perceptual inconsistencies. In contrast, the object-like avatar drew minimal facial attention and prompted participants to shift more quickly toward task-relevant textual information, a pattern consistent with models of attention allocation during low-social-content tasks. These findings reinforce the theoretical view that the visual design of virtual agents influences not only affective impressions but also fundamental attentional strategies, which can alter the depth and quality of user engagement.

Taken together, the multimodal findings show that virtual doctor appearance shapes user experience through interconnected emotional, cognitive, and physiological pathways. The results extend CASA and uncanny valley frameworks by demonstrating that emotional safety, affiliative design cues, and perceptual fluency are critical moderators of human–agent interaction, particularly in mental-health settings. From an applied perspective, the consistent advantages of the animal-like avatar suggest that designs emphasizing warmth, low threat, and socio-emotional comfort may be especially suitable for digital counseling or early intervention contexts. Human-like avatars remain promising but require advances in expressive fidelity and naturalistic micro-behaviors to avoid subtle perceptual inconsistencies. Object-like avatars, while simple to implement, appear limited in their ability to foster therapeutic engagement.

Clinically, these findings provide actionable insights for the design of virtual mental-health systems. Animal-like embodiments may help reduce initial anxiety, increase willingness to disclose personal information, and support emotional regulation, making them well suited for triage systems, youth mental-health services, or supportive check-ins. Human-like embodiments may be preferred for psychoeducation or structured guidance, if realism is sufficiently high to avoid perceptual anomalies. Furthermore, incorporating adaptive avatar systems capable of adjusting visual cues based on real-time physiological or attentional feedback represents a promising direction for personalized digital therapeutic tools. Future work should examine how these design principles generalize across cultures, age groups, and clinical populations, and whether long-term exposure to different avatar types influences therapeutic outcomes, adherence, or emotional resilience.

## 5. Conclusions

The present study examined how three types of virtual doctor embodiments influence users' emotional reactions, physiological regulation, user experience, and visual attention during mental health interactions. By integrating subjective reports, physiological signals, and gaze behavior, the findings demonstrate that avatar appearance exerts measurable effects on both experiential and behavioral outcomes in digital therapeutic settings.

Across multiple indicators, the animal-like avatar produced the most favorable interaction outcomes. It was associated with more positive user experience evaluations, stronger autonomic regulation, and greater social attention. The human-like avatar elicited moderate levels of engagement but did not show strong signs of uncanny valley effects. The object-like avatar consistently produced the lowest levels of emotional and social involvement. These patterns indicate that avatar design influences user comfort and engagement and that effective virtual clinicians should align with users' preferences for emotional safety and social clarity.

The results also underscore the value of incorporating multimodal measurements when evaluating digital mental health systems. Physiological indicators and visual attention patterns were particularly sensitive to differences between avatar types, highlighting the importance of integrating objective behavioral measures alongside self-report data when assessing emotional and cognitive engagement.

Future research should investigate how avatar appearance interacts with verbal style, emotional expressiveness, and dynamic behavioral cues. Additional studies are needed to explore long-term effects on therapeutic alliance, emotional regulation, and treatment adherence. Furthermore, adaptive avatar systems may be developed to tailor visual embodiment to individual user profiles or psychological needs. Together, these directions will support the development of virtual clinicians that are both emotionally supportive and effective in mental health interventions.

**Author Contributions:** Conceptualization, H.Z.; methodology, H.Z.; software, H.Z.; validation, H.Z.; formal analysis, H.Z. and S.W.; investigation, R.P.; resources, H.Z. and R.P.; data curation, H.Z. and R.P.; writing—original draft preparation, H.Z., S.W. and R.P.; writing—review and editing, H.Z.; visualization, S.W.;

supervision, H.Z.; project administration, H.Z.; funding acquisition, H.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the MOE (Ministry of Education in China) Project of Humanities and Social Sciences, grant number 22YJC760044.

**Institutional Review Board Statement:** The study was conducted according to the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board of CENTRAL CHINA NORMAL UNIVERSITY (protocol code CCNU-IRB-202306002).

**Informed Consent Statement:** All participants received written information about the research project, benefits and risks of participation. They were informed that they could withdraw from the study at any time. Written informed consent was obtained prior to the experiment.

**Data Availability Statement:** The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author(s).

**Conflicts of Interest:** The authors declare no conflicts of interest.

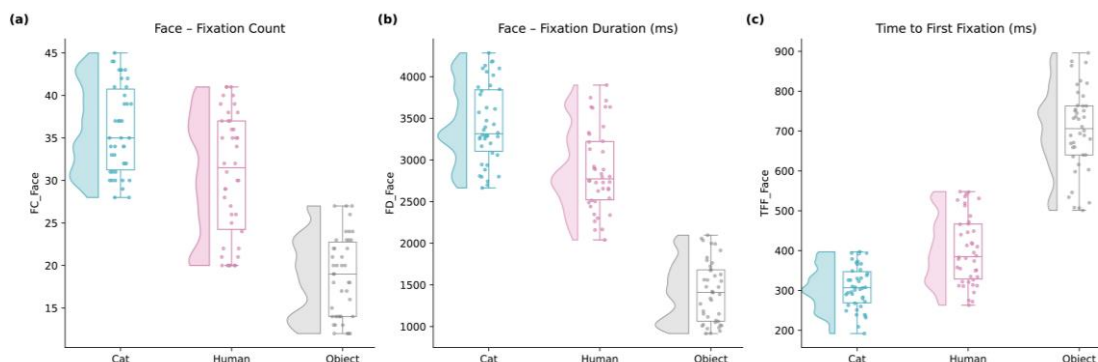
## Abbreviations

The following abbreviations are used in this manuscript:

CASA	Computers as Social Actors
HCI	Human Computer Interaction
HRV	Heart Rate Variability

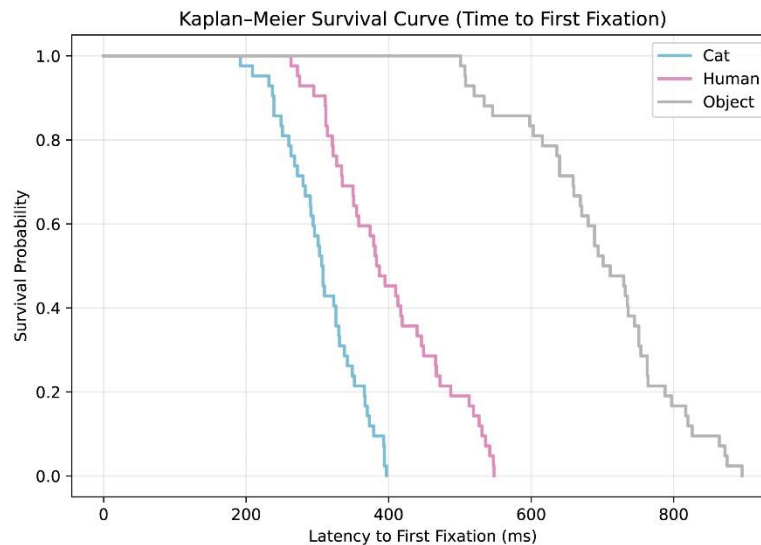
## Appendix A

### Appendix A.1



**Figure A1.** Distribution of Face AOI attention across avatar conditions. (a-c) Raincloud plots showing the distribution, kernel density estimates, and individual data points for fixation count, total fixation duration, and time to first fixation on the avatar's face. Cat, Human, and Object avatar conditions are visualized using violin, box, and jittered point components. All data represent repeated-measures observations.

### Appendix A.2



**Figure A2.** Kaplan–Meier survival curves for time to first fixation on the avatar’s face. Survival curves represent the probability that participants had not yet fixated the avatar’s face over time during the interaction task. Separate curves are shown for Cat, Human, and Object avatars. Curves were estimated using the Kaplan–Meier method, with shaded bands indicating 95% confidence intervals. A steeper decline indicates faster initial visual engagement.

## References

- Ahmed, E., Buruk, O. ‘Oz,’ & Hamari, J. (2024). Human–Robot Companionship: Current Trends and Future Agenda. *International Journal of Social Robotics*, 16(8), 1809–1860. <https://doi.org/10.1007/s12369-024-01160-y>
- Artemiou, E., Hutchison, P., Machado, M., Ellis, D., Bradtke, J., Pereira, M. M., Carter, J., & Bergfelt, D. (2021). Impact of Human–Animal Interactions on Psychological and Physiological Factors Associated With Veterinary School Students and Donkeys. *Frontiers in Veterinary Science*, 8, 701302. <https://doi.org/10.3389/fvets.2021.701302>
- Baake, J., Schmitt, J., & Metag, J. (2025). Balancing Realism and Trust: AI Avatars In Science Communication. *Journal of Science Communication*, 24(2). <https://doi.org/10.22323/2.24020203>
- Beetz, A., Uvnäs-Moberg, K., Julius, H., & Kotrschal, K. (2012). Psychosocial and Psychophysiological Effects of Human-Animal Interactions: The Possible Role of Oxytocin. *Frontiers in Psychology*, 3. <https://doi.org/10.3389/fpsyg.2012.00234>
- Bhattacharya, N., & Gwizdka, J. (2018). Relating Eye-Tracking Measures With Changes In Knowledge on Search Tasks. <https://doi.org/10.48550/ARXIV.1805.02399>
- Bilalpur, M., Inan, M., Zeinali, D., Cohn, J. F., & Alikhani, M. (2024). Learning to Generate Context-Sensitive Backchannel Smiles for Embodied AI Agents with Applications in Mental Health Dialogues (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2402.08837>
- Capozzi, F., & Kingstone, A. (2024). The effects of visual attention on social behavior. *Social and Personality Psychology Compass*, 18(1), e12910. <https://doi.org/10.1111/spc3.12910>
- Cheetham, M. (2011). The human likeness dimension of the “uncanny valley hypothesis”: Behavioral and functional MRI findings. *Frontiers in Human Neuroscience*, 5. <https://doi.org/10.3389/fnhum.2011.00126>
- Cheng, J. T., Gerpott, F. H., Benson, A. J., Bucker, B., Foulsham, T., Lansu, T. A. M., Schülke, O., & Tsuchiya, K. (2023). Eye gaze and visual attention as a window into leadership and followership: A review of empirical insights and future directions. *The Leadership Quarterly*, 34(6), 101654. <https://doi.org/10.1016/j.leaqua.2022.101654>

- Cho, Y. M., Rai, S., Ungar, L., Sedoc, J., & Guntuku, S. C. (2023). An Integrative Survey on Mental Health Conversational Agents to Bridge Computer Science and Medical Perspectives (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2310.17017>
- Cuadrado, F., Lopez-Cobo, I., Mateos-Blanco, T., & Tajadura-Jiménez, A. (2020). Arousing the Sound: A Field Study on the Emotional Impact on Children of Arousing Sound Design and 3D Audio Spatialization in an Audio Story. *Frontiers in Psychology*, 11, 737. <https://doi.org/10.3389/fpsyg.2020.00737>
- De Borst, A. W., & De Gelder, B. (2015). Is it the real deal? Perception of virtual characters versus humans: an affective cognitive neuroscience perspective. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00576>
- Dubois-Sage, M., Jacquet, B., Jamet, F., & Baratgin, J. (2023). We Do Not Anthropomorphize a Robot Based Only on Its Cover: Context Matters too! *Applied Sciences*, 13(15), 8743. <https://doi.org/10.3390/app13158743>
- End, A., & Gamer, M. (2017). Preferential Processing of Social Features and Their Interplay with Physical Saliency in Complex Naturalistic Scenes. *Frontiers in Psychology*, 8. <https://doi.org/10.3389/fpsyg.2017.00418>
- Fraser, A., Hollett, R., Speelman, C., & Rogers, S. L. (2025). Behavioural Realism and Its Impact on Virtual Reality Social Interactions Involving Self-Disclosure. *Applied Sciences*, 15(6), 2896. <https://doi.org/10.3390/app15062896>
- Gullett, N., Zajkowska, Z., Walsh, A., Harper, R., & Mondelli, V. (2023). Heart rate variability (HRV) as a way to understand associations between the autonomic nervous system (ANS) and affective states: A critical review of the literature. *International Journal of Psychophysiology*, 192, 35–42. <https://doi.org/10.1016/j.ijpsycho.2023.08.001>
- Guo, Z., Lai, A., Thygesen, J. H., Farrington, J., Keen, T., & Li, K. (2024). Large Language Model for Mental Health: A Systematic Review. <https://doi.org/10.48550/ARXIV.2403.15401>
- Haensch, A.-C. (2025). "It Listens Better Than My Therapist": Exploring Social Media Discourse on LLMs as Mental Health Tool (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2504.12337>
- Ho, A., Hancock, J., & Miner, A. S. (2018). Psychological, Relational, and Emotional Effects of Self-Disclosure After Conversations With a Chatbot. *Journal of Communication*, 68(4), 712–733. <https://doi.org/10.1093/joc/jqy026>
- Huang, J., Huang, M., Zhan, M., & Guan, D. (2025). The effect of the realism degree of avatars in social virtual worlds: The perspective of self-presentation. *Information & Management*, 62(7), 104185. <https://doi.org/10.1016/j.im.2025.104185>
- Ichikawa, A., Ihara, K., & Kawaguchi, I. (2023). Investigation of How Animal Avatar Affects Users' Self-Disclosure and Subjective Responses in One-on-One Interactions in VR Space. *Asian HCI Symposium'23*, 70–75. <https://doi.org/10.1145/3604571.3604582>
- Itier, R. J., & Batty, M. (2009). Neural bases of eye and gaze processing: The core of social cognition. *Neuroscience & Biobehavioral Reviews*, 33(6), 843–863. <https://doi.org/10.1016/j.neubiorev.2009.02.004>
- Jang, B., Yuh, C., Lee, H., Shin, Y.-B., Lee, H.-J., Kang, E. K., Heo, J., & Cho, C.-H. (2025). Exploring User Experience and the Therapeutic Relationship of Short-Term Avatar-Based Psychotherapy: Qualitative Pilot Study. *JMIR Human Factors*, 12, e66158–e66158. <https://doi.org/10.2196/66158>
- Kang, H. (2021). Sample size determination and power analysis using the G\*Power software. *Journal of Educational Evaluation for Health Professions*, 18, 17. <https://doi.org/10.3352/jeehp.2021.18.17>
- Kawakami, K., Meyers, C., & Fang, X. (2024). Social Cognition, Attention, and Eye Tracking. In D. E. Carlston, K. Hugenberg, & K. L. Johnson (Eds.), *The Oxford Handbook of Social Cognition*, Second Edition (2nd ed., pp. 143–170). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780197763414.013.5>
- Kim, H.-G., Cheon, E.-J., Bai, D.-S., Lee, Y. H., & Koo, B.-H. (2018). Stress and Heart Rate Variability: A Meta-Analysis and Review of the Literature. *Psychiatry Investigation*, 15(3), 235–245. <https://doi.org/10.30773/pi.2017.08.17>
- Klein, S. H. (2025). The effects of human-like social cues on social responses towards text-based conversational agents—A meta-analysis. *Humanit Soc Sci Commun*, 12(1322). <https://doi.org/10.1057/s41599-025-05618-w>

- Kohout, S., Kruikemeier, S., & Bakker, B. N. (2023). May I have your Attention, please? An eye tracking study on emotional social media comments. *Computers in Human Behavior*, 139, 107495. <https://doi.org/10.1016/j.chb.2022.107495>
- Laborde, S., Mosley, E., & Thayer, J. F. (2017). Heart Rate Variability and Cardiac Vagal Tone in Psychophysiological Research – Recommendations for Experiment Planning, Data Analysis, and Data Reporting. *Frontiers in Psychology*, 08. <https://doi.org/10.3389/fpsyg.2017.00213>
- Lawrance, E. L., Thompson, R., Newberry Le Vay, J., Page, L., & Jennings, N. (2022). The Impact of Climate Change on Mental Health and Emotional Wellbeing: A Narrative Review of Current Evidence, and its Implications. *International Review of Psychiatry*, 34(5), 443–498. <https://doi.org/10.1080/09540261.2022.2128725>
- Lawrence, H. R., Schneider, R. A., Rubin, S. B., Mataric, M. J., McDuff, D. J., & Bell, M. J. (2024). The opportunities and risks of large language models in mental health. <https://doi.org/10.48550/ARXIV.2403.14814>
- Lim, S., Schmäzle, R., & Bente, G. (2024). Artificial social influence via human-embodied AI agent interaction in immersive virtual reality (VR): Effects of similarity-matching during health conversations (Version 2). arXiv. <https://doi.org/10.48550/ARXIV.2406.05486>
- Liu, J.-D., You, R.-H., Liu, H., & Chung, P.-K. (2020). Chinese version of the international positive and negative affect schedule short form: Factor structure and measurement invariance. *Health and Quality of Life Outcomes*, 18(1), 285. <https://doi.org/10.1186/s12955-020-01526-6>
- Llanes-Jurado, J., Gómez-Zaragoza, L., Minissi, M. E., Alcañiz, M., & Marín-Morales, J. (2024). Developing conversational Virtual Humans for social emotion elicitation based on large language models. *Expert Systems with Applications*, 246, 123261. <https://doi.org/10.1016/j.eswa.2024.123261>
- Mauss, I. B., & Robinson, M. D. (2009). Measures of emotion: A review. *Cognition & Emotion*, 23(2), 209–237. <https://doi.org/10.1080/02699930802204677>
- Mazerant, K., Van Berlo, Z. M. C., Schouten, A. P., & Willemsen, L. M. (2025). Human nature in a virtual world: The attribution of mind perception to avatars. *Computers in Human Behavior: Artificial Humans*, 6, 100222. <https://doi.org/10.1016/j.chbah.2025.100222>
- Mccraty, R., & Shaffer, F. (2015). Heart Rate Variability: New Perspectives on Physiological Mechanisms, Assessment of Self-regulatory Capacity, and Health Risk. *Global Advances in Health and Medicine*, 4(1), 46–61. <https://doi.org/10.7453/gahmj.2014.073>
- Mohammed, H. (2023). Technology in Association With Mental Health: Meta-ethnography (Version 2). arXiv. <https://doi.org/10.48550/ARXIV.2307.10513>
- Mori, M., MacDorman, K., & Kageki, N. (2012). The Uncanny Valley [From the Field]. *IEEE Robotics & Automation Magazine*, 19(2), 98–100. <https://doi.org/10.1109/MRA.2012.2192811>
- Nass, C., Steuer, J., & Tauber, E. R. (1994). Computers are social actors. *Conference Companion on Human Factors in Computing Systems - CHI '94*, 204. <https://doi.org/10.1145/259963.260288>
- Nowak, K. L., & Biocca, F. (2003). The Effect of the Agency and Anthropomorphism on Users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 12(5), 481–494. <https://doi.org/10.1162/105474603322761289>
- Oh, C. S., Bailenson, J. N., & Welch, G. F. (2018). A Systematic Review of Social Presence: Definition, Antecedents, and Implications. *Frontiers in Robotics and AI*, 5, 114. <https://doi.org/10.3389/frobt.2018.00114>
- Premathilake, G. W., & Li, H. (2024). Users' responses to humanoid social robots: A social response view. *Telematics and Informatics*, 91, 102146. <https://doi.org/10.1016/j.tele.2024.102146>
- Rehm, I. C., Foenander, E., Wallace, K., Abbott, J.-A. M., Kyrios, M., & Thomas, N. (2016). What Role Can Avatars Play in e-Mental Health Interventions? Exploring New Models of Client–Therapist Interaction. *Frontiers in Psychiatry*, 7. <https://doi.org/10.3389/fpsyg.2016.00186>
- Reuben, A., Manczak, E. M., Cabrera, L. Y., Alegria, M., Bucher, M. L., Freeman, E. C., Miller, G. W., Solomon, G. M., & Perry, M. J. (2022). The Interplay of Environmental Exposures and Mental Health: Setting an Agenda. *Environmental Health Perspectives*, 130(2), 025001. <https://doi.org/10.1289/EHP9889>
- Riek, L. D. (2015). Robotics Technology in Mental Health Care. <https://doi.org/10.48550/ARXIV.1511.02281>
- Sakhrani, H., Parekh, S., & Mahajan, S. (2021). Coral: An Approach for Conversational Agents in Mental Health Applications (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2111.08545>

- Schneider, A. K. E., & Bubeck, M. J. (2025). The Dual Character of Animal-Centred Care: Relational Approaches in Veterinary and Animal Sanctuary Work. *Veterinary Sciences*, 12(8), 696. <https://doi.org/10.3390/vetsci12080696>
- Schneider, S., Beege, M., Nebel, S., Schnaubert, L., & Rey, G. D. (2022). The Cognitive-Affective-Social Theory of Learning in digital Environments (CASTLE). *Educational Psychology Review*, 34(1), 1–38. <https://doi.org/10.1007/s10648-021-09626-5>
- Schuetzler, R. M., Giboney, J. S., Grimes, G. M., & Nunamaker, J. F. (2018). The influence of conversational agent embodiment and conversational relevance on socially desirable responding. *Decision Support Systems*, 114, 94–102. <https://doi.org/10.1016/j.dss.2018.08.011>
- Sestino, A., & D'Angelo, A. (2023). My doctor is an avatar! The effect of anthropomorphism and emotional receptivity on individuals' intention to use digital-based healthcare services. *Technological Forecasting and Social Change*, 191, 122505. <https://doi.org/10.1016/j.techfore.2023.122505>
- Shaffer, F., & Ginsberg, J. P. (2017). An Overview of Heart Rate Variability Metrics and Norms. *Frontiers in Public Health*, 5, 258. <https://doi.org/10.3389/fpubh.2017.00258>
- Sinatra, A. M., Pollard, K. A., Files, B. T., Oiknine, A. H., Ericson, M., & Khooshabeh, P. (2021). Social fidelity in virtual agents: Impacts on presence and learning. *Computers in Human Behavior*, 114, 106562. <https://doi.org/10.1016/j.chb.2020.106562>
- Stephenson, M. D., Thompson, A. G., Merrigan, J. J., Stone, J. D., & Hagen, J. A. (2021). Applying Heart Rate Variability to Monitor Health and Performance in Tactical Personnel: A Narrative Review. *International Journal of Environmental Research and Public Health*, 18(15), 8143. <https://doi.org/10.3390/ijerph18158143>
- Tarvainen, M. P., Niskanen, J.-P., Lipponen, J. A., Ranta-aho, P. O., & Karjalainen, P. A. (2014). Kubios HRV – Heart rate variability analysis software. *Computer Methods and Programs in Biomedicine*, 113(1), 210–220. <https://doi.org/10.1016/j.cmpb.2013.07.024>
- Thaler, M., Schlögl, S., & Groth, A. (2021). Agent vs. Avatar: Comparing Embodied Conversational Agents Concerning Characteristics of the Uncanny Valley. <https://doi.org/10.48550/ARXIV.2104.11043>
- Tsang, V. (2018). Eye-tracking study on facial emotion recognition tasks in individuals with high-functioning autism spectrum disorders. *Autism*, 22(2), 161–170. <https://doi.org/10.1177/1362361316667830>
- Wang, Z., Yuan, F., LeBaron, V., Flickinger, T., & Barnes, L. E. (2024). PALLM: Evaluating and Enhancing PALLiative Care Conversations with Large Language Models (Version 2). [arXiv. https://doi.org/10.48550/ARXIV.2409.15188](https://doi.org/10.48550/ARXIV.2409.15188)
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063–1070. <https://doi.org/10.1037/0022-3514.54.6.1063>
- Werkmeister, B., Haase, A. M., Fleming, T., & Officer, T. N. (2024). Environmental Factors for Sustained Telehealth Use in Mental Health Services: A Mixed Methods Analysis. *International Journal of Telemedicine and Applications*, 2024(1), 8835933. <https://doi.org/10.1155/2024/8835933>
- Xu, K., Chen, X., & Huang, L. (2022). Deep mind in social responses to technologies: A new approach to explaining the Computers are Social Actors phenomena. *Computers in Human Behavior*, 134, 107321. <https://doi.org/10.1016/j.chb.2022.107321>

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.