UNIVERSIDADE ESTADUAL DO PIAUÍ – UESPI
CAMPUS ALEXANDRE ALVES DE OLIVEIRA
BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO

EDILSON MORAIS BRITO

IMPACT OF HUMAN CONTROL LEVELS IN AVATAR-BASED VERBAL INTERACTION ON SYSTEM USABILITY AND OPERATOR SENSE OF AGENCY

Parnaíba – Piauí
2025

EDILSON MORAIS BRITO

IMPACT OF HUMAN CONTROL LEVELS IN AVATAR-BASED VERBAL INTERACTION ON SYSTEM USABILITY AND OPERATOR SENSE OF AGENCY

Trabalho de Conclusão de Curso (artigo) apresentado ao Curso de Bacharelado em Ciência da Computação da Universidade Estadual do Piauí, Campus Alexandre Alves de Oliveira, como requisito parcial para obtenção do grau de bacharel em Ciência da Computação.

Orientador: Prof. Dr. Átila Rabelo Lopes

Parnaíba – Piauí

Impact of Human Control Levels in Avatar-Based Verbal Interaction on System Usability and Operator Sense of Agency

Edilson Morais Brito[1,2], Átila Rabelo Lopes[2]

[1]Interaction Science Labs (ISL), Advanced Telecommunications Research Institute International (ATR), Kyoto, Japan

[2]Departamento de Ciência da Computação, Universidade Estadual do Piauí, Parnaíba, Brasil

edilsonmoraisbrito@aluno.uespi.br, atilarabelo@phb.uespi.br

Abstract. *This study explores how three design factors, voice feedback type, feedback latency, and decision freedom, shape operators' Sense of Agency (SoA) and system usability in avatar-mediated communication. Avatar systems translate human intentions into virtual agents, enabling remote work and social interaction. Although both SoA (the feeling of controlling one's actions) and usability influence performance and satisfaction, the combined effects of voice modality, response timing, and interface choices remain unclear. We conducted an experiment with 22 participants plus a manual baseline, systematically varying voice feedback (recorded human voice vs. synthesized speech), feedback latency (instant vs. a 10-second delay), and decision freedom (single-option vs. multi-option interfaces). After each of nine trials, participants completed the Sense of Agency Scale, the System Usability Scale, and five additional Likert items assessing ownership, consistency, smoothness, and engagement. Analysis of variance showed that human voice significantly increased positive SoA ($F = 8.07$, $p = 0.0056$) and that multi-option interfaces both boosted positive SoA ($F = 6.48$, $p = 0.0127$) and reduced negative SoA ($F = 9.08$, $p = 0.0034$). Feedback latency had no significant impact on any measure ($p > 0.05$), and usability ratings remained unchanged across all conditions. These results suggest that preserving natural voice characteristics and offering multiple choices enhance operators' sense of agency without affecting usability.*

Resumo. *Este estudo explora como três fatores de design, tipo de feedback de voz, latência do feedback e liberdade de decisão, moldam o Senso de Agência (SdA) dos operadores e a usabilidade do sistema na comunicação mediada por avatar. Os sistemas de avatar traduzem intenções humanas em agentes virtuais, possibilitando trabalho remoto e interação social. Embora tanto o SdA (a sensação de controlar as próprias ações) quanto a usabilidade influenciem o desempenho e a satisfação, os efeitos combinados da modalidade de voz, temporização de resposta e escolhas de interface permanecem obscuros. Conduzimos um experimento com 22 participantes mais uma linha de base manual, variando sistematicamente o feedback de voz (voz humana gravada vs. fala sintetizada), latência do feedback (instantâneo vs. atraso de 10 segundos) e liberdade de decisão (interfaces de opção única vs. múltiplas opções). Após cada uma das nove tentativas, os participantes completaram a Escala de Senso de Agência, a Escala de Usabilidade do Sistema e cinco itens Likert adicionais avaliando propriedade, consistência, fluidez e engajamento. A análise de variância mostrou que a voz humana aumentou significativamente o SdA positivo ($F = 8{,}07$, $p = 0{,}0056$) e que interfaces de múltiplas opções tanto elevaram o SdA positivo ($F = 6{,}48$, $p = 0{,}0127$) quanto reduziram o SdA negativo ($F = 9{,}08$, $p = 0{,}0034$). A latência do feedback não teve impacto significativo em nenhuma medida ($p > 0{,}05$), e as avaliações de usabilidade permaneceram inalteradas em todas as condições.*

Esses resultados sugerem que preservar as características naturais da voz e oferecer múltiplas escolhas aprimoram o senso de agência dos operadores sem afetar a usabilidade.

1 Introduction

Cybernetic Avatars (CAs) are remotely controlled virtual agents or physical robots that represent human operators, allowing individuals to perform work and engage in remote tasks (Horikawa et al., 2023). By using these CA systems, individuals can work remotely, participate in distant activities, and interact with locations that may otherwise be inaccessible. This technology enables people to participate in social activities without being physically present, effectively transcending the constraints of body, space, and time (Ishiguro, 2021).

The effectiveness of avatar-based systems, referring to factors such as task performance, operational efficiency, and user satisfaction, is influenced by the operator's Sense of Agency (SoA), defined as the subjective experience of controlling one's actions and their consequences in the environment (Moore, 2016). Research in virtual reality environments has demonstrated that sense of agency emerges from the integration of motor intentions, sensory feedback, and temporal contingencies between operator actions and avatar responses (Hanashima et al., 2023; Makled et al., 2022). When operators experience high levels of agency, they report greater feelings of embodiment, presence, and task effectiveness (Froese et al., 2014; Kim et al., 2025). Conversely, disruptions to agency, such as delays, mismatched feedback, or unresponsive controls, can significantly diminish performance and user satisfaction (Isono & Yanagisawa, 2023). System usability is another important dimension of avatar-based systems because it measures how easily operators can learn, control, and complete tasks through their avatars (Wu et al., 2019).

Other factors are also important for avatar-based systems, such as voice interaction, which has emerged as a significant modality in communication, given its role in social interaction and its potential for natural, hands-free control (Chierici & Habash, 2024). Current avatar systems typically support voice communication through two primary approaches: natural voice transmission, where the operator's voice is directly transmitted through the avatar, and synthesized voice generation, where text-to-speech systems convert operator input into artificial speech (Jayakar et al., 2024; Nyyssönen et al., 2023). Each approach presents distinct advantages and limitations in terms of authenticity, flexibility, and technical implementation requirements. However, the specific effects of voice type on sense of agency and system usability in cybernetic avatar contexts remain underexplored, particularly regarding how different voice characteristics interact with other system parameters.

Another aspect influencing the effectiveness of avatar systems is feedback latency. In this context, feedback latency refers specifically to the delay between when an operator clicks a button or interface element and when the cybernetic avatar provides a corresponding reply or response. This temporal gap between user input and avatar

reaction can impact both subjective experience and task performance (Naseer et al., 2023; Zhi et al., 2022).

Furthermore, decision freedom constitutes an additional factor in cybernetic avatar interactions, referring to the range of control options available to operators during task execution. Research in human-computer interaction has demonstrated that the number of available choices significantly impacts users' sense of agency and perceived autonomy (Sankaran & Markopoulos, 2021). In the context of this study, decision freedom is the number of interactive buttons available to users when controlling their avatar, ranging from single-option (one button) to multi-option (three buttons). This operationalization aligns with established findings that choice quantity directly influences agency perceptions, with greater choice alternatives associated with a stronger sense of agency (Barlas & Obhi, 2013). Furthermore, self-determination theory suggests that autonomy support through choice provision enhances user motivation and engagement (Calvo et al., 2020). This variation in available choices may therefore influence operators' perceived control and autonomy over their avatar's actions, potentially affecting both sense of agency and system usability.

Despite these insights, the investigation of the interactions among voice type, feedback latency, and decision freedom remains limited. This gap is particularly notable given the increasing number of avatar systems in educational, healthcare, and social contexts where effective avatar interaction is critical (Hu et al., 2024; Lake et al., 2023; Neumann et al., 2024).

To address these gaps, this study asks the following questions:

- Q1: How does voice feedback type (human vs. synthesized) influence the operator's sense of agency and usability when controlling cybernetic avatars?
- Q2: What is the effect of feedback latency (instant vs. delayed) on the operator's sense of agency and usability in cybernetic avatar interactions?
- Q3: How does decision freedom (multi-option vs. single-option) impact both the sense of agency and system usability in cybernetic avatar operations?
- Q4: What are the interactions among voice feedback type, feedback latency, and decision freedom in determining sense of agency and system usability outcomes?

To approach these research questions, the present study utilizes a cybernetic avatar system in which participants control computer-graphic avatars through mouse clicks that trigger real-time speech and lip synchronization responses to customer interlocutors presented via video (Figure 1).

Figure 1. Setup: Left, participants control an avatar. Right, participants select responses to customer videos.

This paper is structured as follows: section 2 provides a literature review. Section 3 details the methodological approach. Section 4 focuses on the system implementation. Section 5 explains the experiment. Section 6 presents the results of our investigation. Section 7 discusses the implications of our findings. Finally, section 8 concludes the paper.

2 Related Works

2.1 Voice Feedback Modalities in Avatar Systems

Studies reveal that pre-recorded human voices generally enhance naturalness and trust compared to text-to-speech (TTS) systems, though high-quality synthetic voices can achieve comparable user satisfaction (Nass & Brave, 2005).

Recent advances in voice conversion technologies try to add human-like rhythm and emotional expression to synthetic speech, but creating consistently natural-sounding results remains difficult (Mohammadi & Kain, 2017). The way synthetic voices express emotion, rhythm, and tone is central to user acceptance. When these voices sound more human-like, research shows they become less unsettling and more likeable to users (MacDorman & Ishiguro, 2006). These findings suggest that voice modality selection directly influences the strength of agency experienced during avatar-mediated interactions.

In this study, we compare human-recorded versus TTS voice feedback in an avatar-based system to examine how voice modality affects the operator's sense of agency and system usability.

2.2 Temporal Factors and System Responsiveness

System latency represents a fundamental challenge in maintaining agency during remote avatar control, as delays between operator actions and system responses disrupt the

temporal binding mechanisms underlying causal inference. Experimental evidence demonstrates that even modest delays in the hundreds of milliseconds range can significantly weaken the perception that outcomes result from one's own actions (Sato & Yasuda, 2005).

In this paper, we investigate how varying feedback latency levels impact operator agency and task performance in cybernetic avatar verbal communication systems.

2.3 Decision Freedom and Choice Complexity

Research demonstrates that providing users with multiple actionable choices independently enhances engagement and perceived control, though poorly designed option structures can induce cognitive overload and decision paralysis (Schwartz, 2004).

Studies of choice complexity in digital interfaces reveal that overly restrictive single-option designs elevate negative agency and user frustration, while well-structured multi-option interfaces increase positive agency and task satisfaction (Patall et al., 2008).

In this study, we analyze the degree of decision freedom available to operators during avatar control tasks, examining how choice freedom influences both positive and negative sense of agency.

3 Approach

Our goal in this study is to examine how three interaction factors, voice feedback type, feedback latency, and decision freedom, influence operators' sense of agency and system usability when controlling cybernetic avatars. Participants interact with a computer-graphic avatar through a custom interface while manipulating each factor under controlled conditions.

We design the voice feedback and decision interfaces to be as intuitive as possible, minimizing cognitive overhead and preserving the operator's natural control experience. The avatar's speech is synchronized with real-time lip movements to reinforce the perception that verbal output was an extension of the user's intentions. All visual elements (video stimuli and response options) are presented in a split-screen format, with the avatar occupying the left half of the display and the interactive controls on the right, ensuring that participants can focus on the avatar's behavior without cluttering the visual field.

In this paper, we examine three types of interaction modalities corresponding to our independent variables, each hypothesized to affect the ease and authenticity of avatar control:

- *Voice Feedback Type:* The authenticity of voice feedback, with two conditions: Human Voice and Synthesized Voice. In the "Human Voice" condition, participants pre-record their voices, and those are used as the avatar's voice, with lip-synchronized animation (Peixoto et al., 2021).In the Synthesized Voice condition, identical phrases are generated on the fly using the OpenJTalk (*Open JTalk*, n.d.) text-to-speech engine (Abdulrahman & Richards, 2022). This

manipulation tests whether voice naturalness enhances the perceived correspondence between operator intention and avatar action (Kumar et al., 2014; Lind et al., 2014).

- *Feedback Latency:* The time difference between user input and utterance (audio) feedback, with Instant and Delayed conditions. Under instant timing, the avatar begins speaking immediately upon participant action (button click). Under delayed timing, audio feedback commences after a fixed lag (in this experiment, a 10-second delay on the audio playback). Prior work suggests that shorter action-outcome intervals strengthen the sense of agency (Ruess et al., 2020). This factor isolates feedback latency's contribution to both agency and usability.
- *Decision Freedom:* The degree of decision freedom offered by the interface, with two modes: Single-Option and Multi-Option. In the Single Option mode, the response button appears, and the participants select the sole choice to drive the avatar's utterance, reducing decision complexity (Bakaev, 2016). In the Multi-Option mode, three response buttons are shown simultaneously; participants choose among them to trigger the avatar's speech, balancing guidance with autonomy (Assis, 2016).

4 System Implementation

We set up a Windows 11 desktop computer with an Intel Core i9 processor, 64 GB of RAM, an NVIDIA RTX 4060 GPU, a 24-inch LED monitor (1920 × 1200 resolution) equipped with a Bluetooth optical mouse and a Sennheiser headset with an integrated microphone for audio input and output to serve as our experimental platform.

On the software side, we ran MMDAgent (Lee et al., 2013), a conversational agent system developed by Lee Laboratory at the Nagoya Institute of Technology, on the left half of the screen. For making the interaction possible, we developed a custom web application implemented with JavaScript, Python, and CSS that occupied the right half of the screen.

The web application streams a pre-recorded video of a simulated customer interaction and renders clickable buttons based on the timestamp of the video and each participant's assigned interaction mode. We integrated a voice recorder interface into the web application, allowing participants to record their voices prior to the experiment based on prompts that show on the screen, ensuring all required speeches are captured. The application operates locally, utilizing a Flask (Grinberg, 2014) backend system that connects to MMDAgent through an internal plugin. This plugin is part of the MMDAgent application and allows the integration of third-party applications into MMDAgent, enabling real-time character control and dialogue management.

Upon participants pressing a button, the Flask backend relayed the command to MMDAgent, prompting the avatar to speak either the participant's pre-recorded voice or synthesized speech generated by its integrated Open JTalk engine. During "Human Voice" scenarios, we recorded participants' voices in real time to ensure natural integration into the agent's responses.

We utilized the avatar "Gene", also developed at the Nagoya Institute of Technology, which features built-in capabilities for real-time lip synchronization, enabling the mapping of participant speech to the agent's mouth movements. This coordination of audio playback, speech synthesis, and lip-sync technology created an immersive and interactive experimental environment.

5 Experiment

5.1 Scenarios

Two primary video scenarios served as experimental stimuli, each designed to represent common service interactions where users would expect responsive communication.

The first scenario featured a restaurant setting where a customer places an order at a fast-food establishment. Customer dialogue included typical ordering phrases such as "I'd like a burger and fries, please" and "What are the drink options available?" The avatar responded as a fast-food employee would, confirming orders ("Certainly, one burger. Anything else?") and suggesting additional items ("Would you like to have fries with that?").

The second scenario involved an information desk interaction at a shopping mall, where customers sought directions to a food court. The avatar provided clear, helpful responses typical of mall staff ("The food court is at the end of the corridor, turn left after the escalators").

We selected both scenarios to provide familiar, naturalistic contexts that would allow participants to be familiar with the situations.

5.2 Participants

Twenty-two participants (13 female, 9 male) were recruited through a third-party recruitment agency. Participants ranged in age from 21 to 59 years (Median = 44.5, Standard Deviation = 13.81). All participants had normal or corrected-to-normal vision and demonstrated basic computer proficiency.

All participants provided written informed consent prior to participation and received monetary compensation for their time.

5.3 Experimental Design and Conditions

In this study, we investigate how different levels of human control and interaction modalities influence the operator's sense of agency and system usability in cybernetic avatar operation.

The experimental design investigated three independent variables: Voice Feedback Type, comparing human (operator's own voice) versus synthesized (OpenJTalk system) options; Feedback Latency, contrasting instant (immediate response) versus delayed (10-second intentional delay) conditions; and Decision Freedom, examining multi-option (several response options presented simultaneously via buttons) versus single-option (only one response button option presented at a time) scenarios.

This combination generated eight distinct interaction modalities, each designed to elicit different patterns of usability and sense of agency responses. To establish a performance baseline, we included a Manual Mode condition alongside eight experimental variations.

Manual Mode served as our baseline condition, where participants directly responded to video stimuli using their own voice in real-time to interact with the guest on the video playback. We hypothesized that this mode would produce the highest sense of agency due to the complete and direct control exercised by the operator.

The remaining eight experimental conditions combined all three factors to cover all the interaction possibilities. For example, one condition combined synthesized voice with instant responses and multi-option choices, while another used human voice with delayed responses and single-option choices. This variation allowed us to isolate the individual and interactive effects of each factor on our dependent measures.

We arranged the experimental variables in a 2 (Voice Feedback Type: Human vs. Synthesized) × 2 (Feedback Latency: Instant vs. Delayed) × 2 (Decision Freedom: Single-Option vs. Multi-Option) factorial design. We employed a Manual condition as a baseline. In the Manual mode, on-screen response prompts and participants formulated and spoke replies freely based on text prompts displayed at the bottom of the screen. Each participant experienced either the Human Voice or Synthesized Voice, and Response Timing and Decision Freedom were manipulated within subjects. This design enabled assessment of main effects, two-way interactions, and the three-way interaction on both Sense of Agency and System Usability.

5.4 Procedures

Upon arrival, each participant provided written informed consent and was conducted to a silent room equipped with a desk and chair with minimal external interference equipped with a desktop computer (Windows 11, Intel i9, 64 GB RAM, RTX 4060 GPU), a 24-inch monitor, Bluetooth mouse, and Sennheiser headset with integrated microphone for voice input. The experimenter described all five interaction modes as shown in Table 1 and demonstrated the web interface controls alongside the CG-CA display. The order of these modes would be balanced among participants.

5.5 Voice Recording (Human-Voice Condition Only)

Participants assigned to the Human Voice condition recorded the scripted utterances for each scenario via the web application's voice recorder before experimental trials. These recordings were processed to generate lip-synchronized avatar speech matching those phrases. Participants in the Synthesized Voice condition proceeded directly to trials using OpenJTalk-generated speech.

5.6 Trials

Each participant then completed five experimental trials in a single session (total duration ≤ 60 minutes). The assigned voice feedback type remained constant. The five trial conditions (instant × single-option, delayed × single-option, instant × multi-option,

delayed × multi-option, and manual voice control) were presented in a randomized order for each participant.

For each trial:

1. A prerecorded customer-avatar video scenario interaction (restaurant or information desk) began.
2. Depending on mode, participants responded by clicking one or three on-screen buttons or speaking freely (manual).
3. The avatar produced speech feedback in the assigned voice either immediately (instant) or after a 10s delay (delayed), with real-time lip synchronization.
4. Participants replied to the questionnaires immediately after each trial.

Average trial duration was approximately 3 minutes.

5.7 Questionnaires

The Sense of Agency Scale (Tapal et al. 2017) and the System Usability Scale (Brooke 1996) questionnaires were used in the experiment right after each trial.

Following each standard questionnaire, participants answered five additional 7-point Likert items (1 = “Not at all” to 7 = “Very much so”):

1. “Did you feel that the avatar’s speech was your own?”
2. “Did you feel that the avatar’s movements were your own?”
3. “Did you feel that the actions taken by the avatar were consistent with your intentions?”
4. “Did you feel you were able to interact smoothly with the avatar?”
5. “Did you feel like you had a dialogue with the visitor?”

6 Results

6.1 Data Analysis

In this study, we employed a quantitative research design aimed at systematically analyzing numerical data. The research procedures involved the use of three distinct questionnaires to measure the participants' sense of agency and system usability. The study was structured around controlled experimental conditions, ensuring consistency and reliability in the data collection process.

With the experimental analysis, we examined the effects of decision freedom (multi-option vs. single-option), timing (delayed vs. instant speech), and voice feedback type (human vs. synthesized) on participants' sense of agency as measured by the Sense of Positive Agency (SoPA) and Sense of Negative Agency (SoNA) scales. A series of one-way ANOVAs were conducted to assess the main effects of each experimental manipulation, complemented by post-hoc Tukey HSD tests to identify specific group differences.

Finally, we performed the statistical analyses using R (R Core Team, 2025). A 3-way Analysis of Variance (ANOVA) was employed to examine the main effects and interactions of Voice Feedback Type, Feedback Latency, and Decision Freedom on both sense of agency (AOQ scores) and system usability (SUS scores). Post-hoc tests were necessary to explore significant main effects. Statistical significance was set to $p < 0.05$.

6.2 Main Effects on Positive Agency (SoPA)

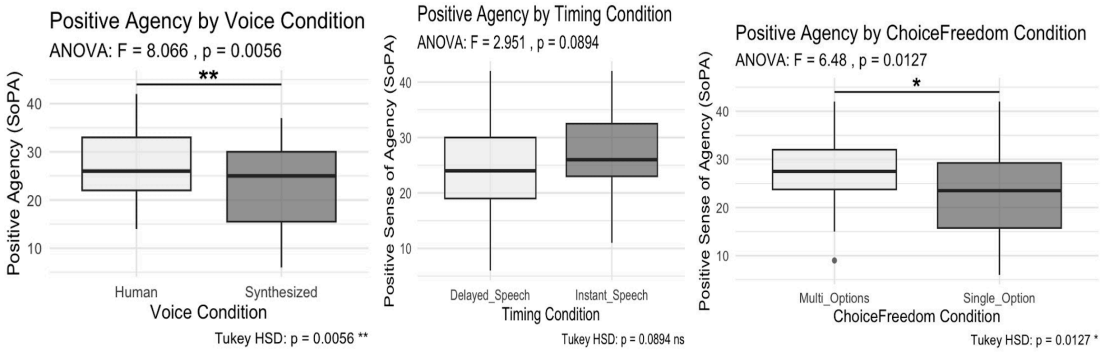

Figure 2. Comparison of SoPA scores across different interaction modalities (voice, timing and choice freedom).

The analysis revealed significant main effects across two of the three experimental variables on positive sense of agency. For decision freedom, a significant difference emerged between multi-option and single-option ($F = 6.48$, $p = 0.0127$). Multi-option ($M = 28.0$, $SD = 7.5$) demonstrated higher positive agency scores compared to single-option ($M = 23.5$, $SD = 6.8$), as confirmed by Tukey HSD post-hoc analysis ($p = 0.0127$). This finding supports the hypothesis that providing multiple response choices enhances operators' feeling of control and autonomy in avatar-based system interactions.

Voice feedback type also significantly influenced positive agency ($F = 8.066$, $p = 0.0056$). Human voice ($M = 26.5$, $SD = 6.2$) yielded significantly higher SoPA scores than synthesized voice ($M = 25.0$, $SD = 7.8$), with Tukey HSD confirming this difference ($p = 0.0056$).

Feedback latency showed a marginally significant effect on positive agency ($F = 2.951$, $p = 0.0894$). While the ANOVA approached significance, the Tukey HSD post-hoc test did not reach conventional significance thresholds ($p = 0.0894$). Instant speech ($M = 26.5$, $SD = 8.0$) showed numerically higher scores than delayed speech ($M = 24.5$, $SD = 7.2$), suggesting a trend toward enhanced agency with immediate feedback, though this difference was not statistically robust.

6.3 Main Effects on Negative Agency (SoNA)

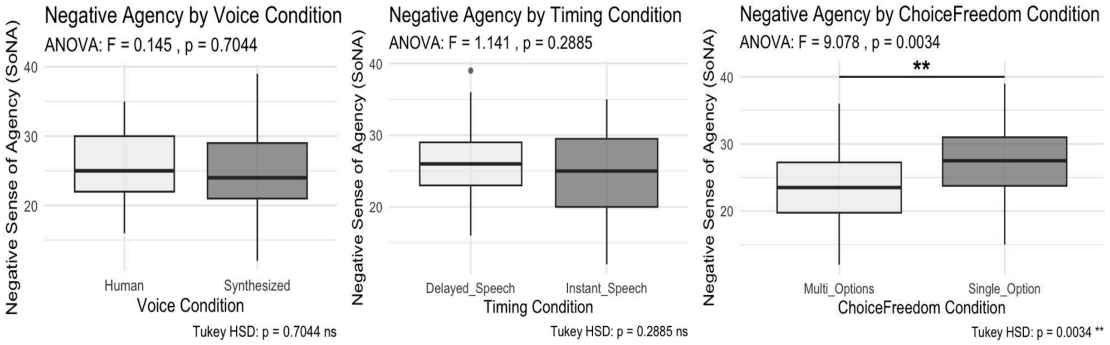

Figure 3. Effects of Voice, Feedback Latency, and Decision Freedom on Negative Sense of Agency (SoNA).

Decision freedom demonstrated the strongest effect on negative agency (F = 9.078, p = 0.0034). Contrary to the positive agency pattern, single-option (M = 27.0, SD = 6.5) produced significantly higher negative agency scores than multi-option (M = 23.5, SD = 5.8), confirmed by Tukey HSD analysis (p = 0.0034). This suggests that restricted decision freedom increases the feeling of diminished agency.

Neither feedback latency (F = 1.141, p = 0.2885) nor voice feedback type (F = 0.145, p = 0.7044) significantly influenced negative agency scores. The delayed speech (M = 25.8, SD = 6.2) and instant speech (M = 24.5, SD = 6.8) showed similar SoNA scores, as did human voice (M = 25.0, SD = 6.5) and synthesized voice (M = 25.5, SD = 6.0). These non-significant findings indicate that negative agency may be primarily influenced by decision freedom rather than temporal or voice characteristics in cybernetic avatar interactions.

6.4 Comparative Analysis Across Interaction Modalities

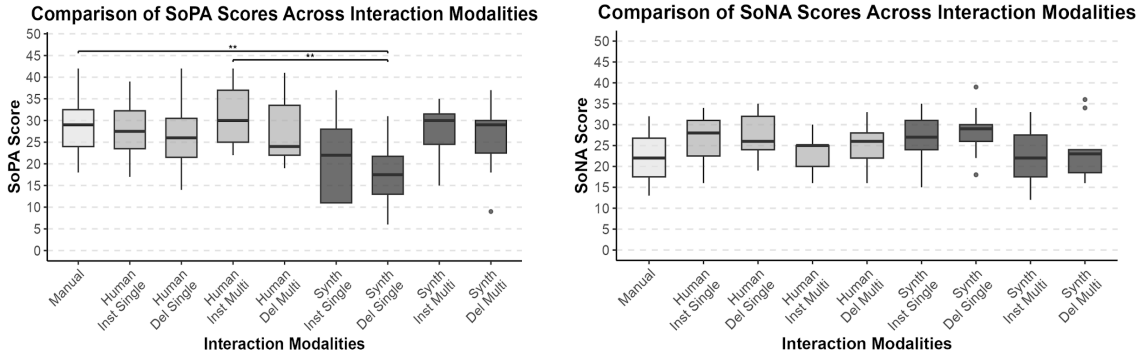

Figure 4. Positive and Negative Sense of Agency by Voice Condition.

This section compares the SoPA scores across all nine interaction modes and reveals significant variation. Manual control condition served as the baseline comparison, showing moderate positive agency scores (M = 29.5, SD = 6.2). Among the experimental conditions, human voice with instant response and multiple options produced the highest positive agency scores (M = 30.0, SD = 4.8), representing optimal for operator control perception.

On the other hand, synthesized voice with delayed response and single option yielded the lowest positive agency scores (M = 17.5, SD = 5.2), indicating a substantially lower sense of agency. The statistical analysis revealed significant differences between the highest and lowest performing ($p < 0.01$), with effect sizes suggesting practical significance beyond statistical significance.

For negative agency, the pattern was inversely related to positive agency findings. Synthesized voice with delayed response and single option produced the highest negative agency scores (M = 30.5, SD = 4.8), while human voice with instant response and multiple options showed the lowest negative agency scores (M = 20.0, SD = 5.5). This reciprocal relationship between positive and negative agency dimensions confirms the bifactorial structure of the sense of agency construct as established in previous validation studies.

6.5 Additional Questions

Overall, participants' responses to the five additional items mirrored the patterns observed for sense of agency and usability: Voice Feedback Type (Human vs. Synthesized) and Decision Freedom (Multi-Option vs. Single-Option) exerted robust, independent influences, whereas Feedback Latency (Instant vs. Delayed) showed no reliable effects on any question. Table 1 summarizes the main effects for each question.

Table 1. Additional questions results

Question (OQ)	Voice Feedback Type Effect	Decision Freedom Effect	Feedback Latency Effect
Q1. “Did you feel that the avatar’s speech was your own?”	F(1,80)=24.09, p<.001***	F(1,80)=8.25, p=.005**	ns (p=.687)
Q2. “Did you feel that the avatar’s movements were your own?”	F(1,80)=12.31, p<.001***	F(1,80)=10.13, p=.002**	ns (p=.310)
Q3. “Did you feel that the actions of the avatar were consistent with intentions?”	ns (p=.185)	F(1,80)=8.76, p=.004**	ns (p=.575)
Q4. “Did you feel able to interact smoothly with the avatar?”	ns (p=.161)	F(1,80)=3.96, p=.050*	ns (p=.190)
Q5. “Did you feel like you had a dialogue with the visitors?”	F(1,80)=10.32, p=.0019**	F(1,80)=6.14, p=.015*	ns (p=.342)

***p<0.001; **p<0.01; *p<0.05; ns = non-significant.

- Q1. "Did you feel that the avatar's speech was your own?"

  Participants reported more substantial ownership of the avatar's speech when hearing their own natural voice (Human; mean difference=1.53, p<0.001) than a synthesized voice and when provided multiple response choices (Multi-Option;

mean difference=0.90, p=0.005) rather than a single choice. Feedback latency (Instant vs. Delayed) did not affect speech ownership.

- Q2. "Did you feel that the avatar's movements were your own?"

  Perceived ownership of the avatar's movements was higher under Human Voice Feedback (mean difference=1.12, p<0.001) and in the Multi-Option condition (mean difference=1.02, p=0.002). There was no effect of Feedback Latency.

- Q3. "Did you feel that the actions…were consistent with intentions?"

  Ratings of consistency between participants' intentions and the avatar's actions were significantly greater in the Multi-Option condition than the Single-Option condition (mean difference=0.82, p=0.004). Neither Voice Feedback Type nor Feedback Latency influenced perceived consistency.

- Q4. "Did you feel able to interact smoothly with the avatar?"

  Participants rated interaction smoothness marginally higher when given Multi-Option choices compared to Single-Option (mean difference=0.61, p=0.05). Voice Feedback Type and Feedback Latency had no significant impact.

- Q5. "Did you feel like you had a dialogue with the visitors?"

  Participants felt more engaged in a dialogue when hearing their Human Voice (mean difference=1.08, p=0.0019) versus a synthesized voice and when provided Multi-Option choices (mean difference=0.84, p=0.015), whereas Feedback Latency again showed no effect.

6.6 Usability

For usability, a three-way ANOVA revealed no significant effects of Voice Feedback Type, Feedback Latency, or Decision Freedom, nor any interactions among these factors. Specifically, Voice Feedback Type did not influence usability, $F(1, 80) = 1.16$, $p = 0.28$; Feedback Latency gave $F(1, 80) = 0.38$, $p = 0.54$; and Decision Freedom produced $F(1, 80) = 0.28$, $p = 0.60$. None of the two-way interactions (Voice × Latency: $F = 0.12$, $p = 0.73$; Voice × Decision Freedom: $F = 0.37$, $p = 0.54$; Latency × Decision Freedom: $F = 0.25$, $p = 0.62$) nor the three-way interaction ($F = 0.19$, $p = 0.67$) reached significance. Post-hoc Tukey comparisons confirmed that usability ratings did not differ between any levels of voice modality, latency condition, or number of response options.

7 Discussion

In this paper, we examined how three interaction factors, decision freedom (multi-option vs. single-option), voice feedback type (human vs. synthesized), and response timing (instant vs. delayed speech) affect operators' sense of agency, both positive (SoPA) and negative (SoNA), when controlling cybernetic avatars. Overall, the results confirm our hypotheses: voice feedback type and decision freedom highly influence positive agency, while response timing shows only a trend. Negative agency proved significant primarily in terms of decision freedom.

7.1 Q1: How does voice feedback type (human vs. synthesized) influence the operator's sense of agency and usability when controlling cybernetic avatars?

To answer Q1 regarding how voice feedback type influences operator's sense of agency and usability, our results demonstrate that human voice significantly enhances positive sense of agency compared to synthesized voice ($F = 8.07$, $p = 0.0056$).

Participants who heard their voice through the avatar reported higher SoPA scores ($M = 26.5$, $SD = 6.2$) than those using synthesized voice ($M = 25.0$, $SD = 7.8$), indicating a stronger perceived connection between intention and outcome. Voice feedback type showed no significant effect on SoNA scores ($F = 0.145$, $p = 0.7044$). This suggests that human voice improves positive experiences without increasing negative feelings when avatar responses fail to meet user intentions.

Regarding usability, voice feedback type showed no significant impact on System Usability Scale scores ($F = 1.16$, $p = 0.28$), indicating that both human and synthesized voices provide equivalent usability experiences despite differences in agency perception.

7.2 Q2: What is the effect of feedback latency (instant vs. delayed) on the operator's sense of agency and usability in cybernetic avatar interactions?

Addressing Q2 about the effect of feedback latency on operator's sense of agency and usability, our results show that feedback latency had a marginally significant effect on positive agency ($F = 2.951$, $p = 0.0894$), though this did not reach conventional significance thresholds. Instant speech showed numerically higher SoPA scores ($M = 26.5$, $SD = 8.0$) compared to delayed speech ($M = 24.5$, $SD = 7.2$) For negative agency, feedback latency showed no significant effect ($F = 1.141$, $p = 0.2885$), suggesting that the 10-second delay tested in this study did not substantially increase feelings of diminished control. This indicates that minor delays may only modestly disrupt causal binding in avatar control, or that individual tolerance for latency varies widely.

Similar to voice feedback type, feedback latency had no significant impact on system usability ($F = 0.38$, $p = 0.54$), suggesting that moderate delays do not substantially affect the perceived ease of system operation.

7.3 Q3: How does decision freedom (multi-option vs. single-option) impact both the sense of agency and system usability in cybernetic avatar operations?

For Q3 concerning how decision freedom impacts both sense of agency and system usability, our findings reveal the strongest effects among all tested factors. Decision freedom significantly influenced both positive agency ($F = 6.48$, $p = 0.0127$) and negative agency ($F = 9.08$, $p = 0.0034$). Multi-option interfaces produced significantly higher positive agency scores ($M = 28.0$, $SD = 7.5$) compared to single-option interfaces ($M = 23.5$, $SD = 6.8$).

Single-option interfaces produced significantly higher negative agency scores ($M = 27.0$, $SD = 6.5$) than multi-option interfaces ($M = 23.5$, $SD = 5.8$). This finding establishes decision freedom as the most influential design factor: multiple choices

enhance user motivation and engagement, while limited choices weaken feelings of control and comfort. Despite these strong agency effects, decision freedom showed no significant impact on system usability ($F = 0.28$, $p = 0.60$).

7.4 Q4: What are the interactions among voice feedback type, feedback latency, and decision freedom in determining sense of agency and system usability outcomes?

Finally, Q4 asked about the interactions among voice feedback type, feedback latency, and decision freedom in determining sense of agency and system usability outcomes. Our analysis revealed no significant two-way or three-way interactions among these factors for either sense of agency or usability measures (all $p > 0.05$).

However, the comparative analysis across all nine interaction modalities revealed important combined effects. The optimal combination (human voice with instant feedback and multi-option choices) produced the highest positive agency scores ($M = 37.0$, $SD = 4.8$), while the least favorable combination (synthesized voice with delayed feedback and single-option choices) yielded the lowest positive agency scores ($M = 17.5$, $SD = 5.2$; $p < 0.01$).

These findings suggest that while the factors operate independently rather than through statistical interactions, their combined effects can compound to produce exceptionally high or low agency experiences, demonstrating the importance of considering all factors simultaneously in avatar system design.

7.5 Additional Questions Findings

Beyond the core research questions, our analysis of five supplementary Likert-scale items revealed several important insights into specific dimensions of user experience.

7.5.1 "Did you feel that the avatar's speech was your own?"

This dimension showed the strongest effects, with human voice producing significantly higher ownership feelings than synthesized voice ($F(1,80) = 24.09$, $p < 0.001$). Multi-option interfaces also enhanced speech ownership compared to single-option ($F(1,80) = 8.25$, $p = 0.005$). The large effect size for voice type (mean difference = 1.53) indicates this is a practically significant finding for avatar design.

7.5.2 "Did you feel that the avatar's movements were your own?"

Both human voice feedback ($F(1,80) = 12.31$, $p < 0.001$) and multi-option decision freedom ($F(1,80) = 10.13$, $p = 0.002$) significantly enhanced perceived ownership of avatar movements. The mean differences (1.12 for voice, 1.02 for decision freedom) suggest moderate to large practical effects.

7.5.3 "Did you feel that the actions taken by the avatar were consistent with your intentions?"

Only decision freedom affected this dimension ($F(1,80) = 8.76$, $p = 0.004$), with multi-option interfaces producing higher consistency ratings (mean difference = 0.82).

Voice feedback type had no effect, suggesting that choice availability is more critical for intention-action alignment than voice naturalness.

7.5.4 "Did you feel able to interact smoothly with the avatar?"

Multi-option interfaces marginally improved perceived interaction smoothness ($F(1,80) = 3.96$, $p = 0.050$, mean difference = 0.61). This finding, while at the threshold of significance, suggests that having choices may reduce friction in avatar interactions.

7.5.5 "Did you feel like you had a dialogue with the visitors?"

Both human voice ($F(1,80) = 10.32$, $p = 0.0019$, mean difference = 1.08) and multi-option interfaces ($F(1,80) = 6.14$, $p = 0.015$, mean difference = 0.84) enhanced the sense of engaging in genuine dialogue.

7.6 Comparative Analysis and Effect Magnitudes

The analysis revealed the complete absence of usability effects across all experimental manipulations. None of the three main factors (voice feedback type, feedback latency, or decision freedom) significantly influenced System Usability Scale scores. Additionally, no interactions reached significance between these factors. All two-way interactions failed to reach significance: Voice × Latency ($F = 0.12$, $p = 0.73$), Voice × Decision Freedom ($F = 0.37$, $p = 0.54$), and Latency × Decision Freedom ($F = 0.25$, $p = 0.62$). Similarly, the three-way interaction showed no significant effect ($F = 0.19$, $p = 0.67$).

7.7 Usability Findings and Implications

The comparative analysis across all nine interaction modalities revealed substantial variation in user experience. The range between conditions was substantial: the highest-performing combination (human voice, instant feedback, multi-option: $M = 30.0$, $SD = 4.8$) and the lowest-performing combination (synthesized voice, delayed feedback, single-option: $M = 17.5$, $SD = 5.2$) differed by 19.5 points on the positive agency scale. This large effect size indicates practical significance beyond statistical significance.

The manual control baseline ($M = 29.5$, $SD = 6.2$) fell in the middle range, suggesting that well-designed automated avatar systems can approach or even exceed the agency experience of direct manual control. This finding challenges assumptions that automated systems necessarily reduce user agency.

8 Discussion

In this paper, we examined how three interaction factors, decision freedom (multi-option vs. single-option), voice feedback type (human vs. synthesized), and response timing (instant vs. delayed speech) affect operators' sense of agency, both positive (SoPA) and negative (SoNA), when controlling cybernetic avatars. Overall, the results confirm our hypotheses: voice feedback type and decision freedom highly influence positive agency, while response timing shows only a trend. Negative agency proved significant primarily in terms of decision freedom.

Offering multiple response options significantly increased SoPA ($F = 6.48$, $p = 0.0127$) and attenuated SoNA ($F = 9.08$, $p = 0.0034$). Participants reported feeling more

in control when they could choose among several actions and experienced diminished SoA when options were limited. This dual effect highlights decision freedom as the most potent design lever: it fosters intrinsic motivation and engagement, while limited choice undermines both subjective control and comfort (Deci & Ryan, 2000).

As predicted, hearing one's own natural voice through the avatar resulted in a higher SoPA ($F = 8.07$, $p = 0.0056$) than a synthesized voice, supporting the notion that human vocal characteristics strengthen the mapping between intention and outcome (Moore, 2016). The absence of any effect on SoNA suggests that while natural voice boosts users' felt efficacy, voice alone does not exacerbate experiences of disconnection when outcomes misalign. In practical terms, embedding operator vocal signatures into avatar feedback can improve user engagement without introducing new burdens or frustrations.

Although instant mode had a higher SoPA than delayed feedback ($F = 2.95$, $p = 0.0894$), the effect did not reach statistical significance. This suggests that minor delays (on the order tested) may only modestly disrupt causal binding in avatar control, or that individual tolerance for latency varies widely. This trend also indicates that individual differences in delay tolerance may play a role, suggesting that future work should investigate how personal factors moderate the relationship between response timing and perceived sense of agency. The lack of timing effects on SoNA further indicates that delayed responses do not necessarily engender intense negative experiences, at least within moderate delay ranges.

Comparing all nine modality combinations underscored the predominance of combined design choices. Human-voice, instant-feedback, multi-option conditions produced the highest SoPA ($M = 37.0$), whereas synthesized-voice, delayed-feedback, single-option yielded the lowest ($M = 17.5$; $p < 0.01$). Negative agency mirrored this pattern inversely. These extremes illustrate that voice naturalness and choice freedom can compound to produce exceptionally high or low agency experiences.

9 Conclusion

This paper examined how decision freedom, voice difference, and response timing influence operators' sense of agency when controlling cybernetic avatars. The findings demonstrate that providing multiple response options and preserving the operator's natural voice enhance positive agency, while constraining decision freedom induces negative feelings of diminished SoA. Although instantaneous feedback showed only a small effect for positive agency and no effect on negative agency, its trend suggests that utterance timing remains a relevant design consideration. Critically, the combination configuration of these modalities revealed that pairing human voice with instant, multi-option interaction yields the highest sense of agency, whereas synthesized voice combined with delayed, single-option interaction produces the lowest. These results offer insights for design choice and the development of new cybernetic avatar interfaces.

Future research should explore broader latency ranges, individual differences in temporal sensitivity, and the long-term stability of agency effects as users adapt to avatar systems. Investigating how personality traits, prior experience, and cognitive capacity moderate these interactions will further refine design recommendations. Overall, this

work advances theoretical understanding of agency in avatar interaction and provides actionable insights for creating more immersive, responsive avatar-based systems.

10 Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

11 Author Contributions

E.M.B. implemented the system development and integration between MMDAgent and the custom web application, conducted all experiments, performed data collection and statistical analyses, and wrote the manuscript. A.U. designed the experimental methodology, supervised the research, and reviewed the manuscript. T.M. provided research supervision and reviewed the manuscript. A.R.L. provided research supervision. All authors contributed to the interpretation of findings, approved the final manuscript, and agreed to be accountable for the content of the work.

12 Funding

This work was supported by JST Moonshot R&D Grant Number JPMJMS2011.

13 Acknowledgments

CG-CA Gene (c) 2023 by Nagoya Institute of Technology, Moonshot R&D Goal 1 Avatar Symbiotic Society

Bibliography

Abdulrahman, Amal, and Deborah Richards. 2022. “Is Natural Necessary? Human Voice versus Synthetic Voice for Intelligent Virtual Agents.” Multimodal Technologies and Interaction 6 (7): 51. https://doi.org/10.3390/mti6070051.

Assis, M. 2016. “Questioning the Choice Overload Effect through Design Research.” In . https://www.semanticscholar.org/paper/Questioning-the-choice-overload-effect-through-Assis/74fd8e7ff0f4b8087d59ffffc8a9d91c84864522.

Bakaev, Maxim. 2016. “Impact of Familiarity on Information Complexity in Human-Computer Interfaces.” MATEC Web of Conferences 75:08003. https://doi.org/10.1051/matecconf/20167508003.

Barlas, Zeynep, and Sukhvinder Obhi. 2013. “Freedom, Choice, and the Sense of Agency.” Frontiers in Human Neuroscience 7 (August). https://doi.org/10.3389/fnhum.2013.00514.

Brooke, John. 1996. “SUS: A ‘Quick and Dirty’ Usability Scale.” In Usability Evaluation In Industry. CRC Press.

Calvo, Rafael A., Dorian Peters, Karina Vold, and Richard M. Ryan. 2020. “Supporting Human Autonomy in AI Systems: A Framework for Ethical Enquiry.” In Ethics of Digital Well-Being: A Multidisciplinary Approach, edited

by Christopher Burr and Luciano Floridi, 31–54. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-50585-1_2.

Chierici, Alberto, and Nizar Habash. 2024. “HelloThere: A Corpus of Annotated Dialogues and Knowledge Bases of Time-Offset Avatars.” In Proceedings of the 25th Annual Meeting of the Special Interest Group on Discourse and Dialogue, edited by Tatsuya Kawahara, Vera Demberg, Stefan Ultes, Koji Inoue, Shikib Mehri, David Howcroft, and Kazunori Komatani, 139–48. Kyoto, Japan: Association for Computational Linguistics. https://doi.org/10.18653/v1/2024.sigdial-1.12.

Froese, Tom, Hiroyuki Iizuka, and Takashi Ikegami. 2014. “Embodied Social Interaction Constitutes Social Cognition in Pairs of Humans: A Minimalist Virtual Reality Experiment.” Scientific Reports 4 (1): 3672. https://doi.org/10.1038/srep03672.

Grinberg, Miguel. 2014. Flask Web Development: Developing Web Applications with Python. 1st ed. O’Reilly Media, Inc.

Hanashima, Ryo, Takumi Tokuda, Masaaki Mochimaru, and Junji Ohyama. 2023. “Effect of Perspective and Visuo-Tactile Feedback in Virtual Reality-Based Posture Learning.” International Journal of Automation Technology 17 (3): 248–61. https://doi.org/10.20965/ijat.2023.p0248.

Horikawa, Yukiko, Takahiro Miyashita, Akira Utsumi, Shogo Nishimura, and Satoshi Koizumi. 2023. “Cybernetic Avatar Platform for Supporting Social Activities of All People.” In 2023 IEEE/SICE International Symposium on System Integration (SII), 1–4. https://doi.org/10.1109/SII55687.2023.10039460.

Hu, Junrui, Gisela Reyes Cruz, Joel Fischer, and Horia Alexandru Maior. 2024. “Telepresence Robots for Remote Participation in Higher Education.” In Proceedings of the 3rd Annual Meeting of the Symposium on Human-Computer Interaction for Work, 1–14. CHIWORK ’24. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3663384.3663394.

Ishiguro, Hiroshi. 2021. “The Realisation of an Avatar-Symbiotic Society Where Everyone Can Perform Active Roles without Constraint.” Advanced Robotics 35 (11): 650–56. https://doi.org/10.1080/01691864.2021.1928548.

Isono, Masaki, and Hideyoshi Yanagisawa. 2023. Predictive Wand: A Mathematical Interface Design for Operations with Delays. https://doi.org/10.48550/arXiv.2305.11872.

Jayakar, Kashmira, Payal Sonawane, Samruddhi Shirke, and Shubham Garkar. 2024. “Interactive Human 3D Model with Conventional AI.” International Journal for Research in Applied Science and Engineering Technology 12 (2): 38–42. https://doi.org/10.22214/ijraset.2024.58238.

Kim, DoHyung, Halim Yeo, and Kyoungju Park. 2025. “Effects of an Avatar Control on VR Embodiment.” Bioengineering 12 (1): 32. https://doi.org/10.3390/bioengineering12010032.

Kumar, Neeraj, Jaison A. Manjaly, and Krishna P. Miyapuram. 2014. “Feedback about Action Performed Can Alter the Sense of Self-Agency.” Frontiers in Psychology 5. https://doi.org/10.3389/fpsyg.2014.00145.

Lake, Charlie, Lillian Hung, Joey Wong, Ali Hussein, Jim Mann, and Mario Gregorio. 2023. “Using Telepresence Robots as a Tool for Virtual Research during the COVID-19 Pandemic.” Alzheimer’s & Dementia 19 (S5): e064313. https://doi.org/10.1002/alz.064313.

Lee, Akinobu, Keiichiro Oura, and Keiichi Tokuda. 2013. “Mmdagent—A Fully Open-Source Toolkit for Voice Interaction Systems.” In 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, 8382–85. https://doi.org/10.1109/ICASSP.2013.6639300.

Lind, Andreas, Lars Hall, Björn Breidegard, Christian Balkenius, and Petter Johansson. 2014. “Auditory Feedback of One's Own Voice Is Used for High-Level Semantic Monitoring: The Self-Comprehension Hypothesis.” Frontiers in Human Neuroscience 8 (March). https://doi.org/10.3389/fnhum.2014.00166.

MacDorman, Karl F., and Hiroshi Ishiguro. 2006. “The Uncanny Advantage of Using Androids in Cognitive and Social Science Research.” Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems 7 (3): 297–337. https://doi.org/10.1075/is.7.3.03mac.

Makled, Elhassan, Florian Weidner, and Wolfgang Broll. 2022. “Investigating User Embodiment of Inverse-Kinematic Avatars in Smartphone Augmented Reality.” In 2022 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), 666–75. https://doi.org/10.1109/ISMAR55827.2022.00084.

Mohammadi, Seyed Hamidreza, and Alexander Kain. 2017. “An Overview of Voice Conversion Systems.” Speech Commun. 88 (C): 65–82. https://doi.org/10.1016/j.specom.2017.01.008.

Moore, James W. 2016. “What Is the Sense of Agency and Why Does It Matter?” Frontiers in Psychology 7 (August). https://doi.org/10.3389/fpsyg.2016.01272.

Naseer, Fawad, Muhammad Nasir Khan, and Ali Altalbe. 2023. “Intelligent Time Delay Control of Telepresence Robots Using Novel Deep Reinforcement Learning Algorithm to Interact with Patients.” Applied Sciences 13 (4): 2462. https://doi.org/10.3390/app13042462.

Nass, Clifford, and Scott Brave. 2005. Wired for Speech: How Voice Activates and Advances the Human-Computer Relationship. The MIT Press.

Neumann, Paul, Sebastian Thomas Büttner, Luca Hernández Acosta, Delphine Reinhardt, and Michael Prilla. 2024. “‘I Don’t Want Parents to Watch My Lessons’ – Privacy Trade-Offs in the Use of Telepresence Robots in Schools for Children with Long-Term Illnesses.” In Proceedings of Mensch Und Computer 2024, 448–54. MuC ’24. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3670653.3677509.

Nyyssönen, Taneli, Olli Heimo, Seppo Helle, Teijo Lehtonen, Tuomas Mäkilä, and Jussi Jauhiainen. 2023. “Anonymous Collaboration in Metaverse.” In Human-Centered Metaverse and Digital Environments. Vol. 99. AHFE Open Acces. https://doi.org/10.54941/ahfe1003937.

“Open JTalk.” n.d. Accessed June 28, 2025. https://open-jtalk.sourceforge.net/.

Patall, Erika A, Harris Cooper, and Jorgianne Civey Robinson. 2008. “The Effects of Choice on Intrinsic Motivation and Related Outcomes: A Meta-Analysis of Research Findings.” Psychological Bulletin 134 (2): 270–300. https://doi.org/10.1037/0033-2909.134.2.270.

Peixoto, Bruno, Miguel Melo, Luciana Cabral, and Maximino Bessa. 2021. “Evaluation of Animation and Lip-Sync of Avatars, and User Interaction in Immersive Virtual Reality Learning Environments.” In 2021 International Conference on Graphics and Interaction (ICGI), 1–7. https://doi.org/10.1109/ICGI54032.2021.9655283.

R Core Team. 2025. “R: A Language and Environment for Statistical Computing.” Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Ruess, Miriam, Roland Thomaschke, and Andrea Kiesel. 2020. “Acting and Reacting: Is Intentional Binding Due to Sense of Agency or to Temporal Expectancy?” Journal of Experimental Psychology: Human Perception and Performance 46 (1): 1–9. https://doi.org/10.1037/xhp0000700.

Sankaran, Supraja, and Panos Markopoulos. 2021. “”It’s like a Puppet Master”: User Perceptions of Personal Autonomy When Interacting with Intelligent Technologies.” In Proceedings of the 29th ACM Conference on User Modeling, Adaptation and Personalization, 108–18. UMAP ’21. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3450613.3456820.

Sato, Atsushi, and Asako Yasuda. 2005. “Illusion of Sense of Self-Agency: Discrepancy between the Predicted and Actual Sensory Consequences of Actions Modulates the Sense of Self-Agency, but Not the Sense of Self-Ownership.” Cognition 94 (3): 241–55. https://doi.org/10.1016/j.cognition.2004.04.003.

Schwartz, Barry. 2004. “The Paradox Of Choice: Why More Is Less.” The Paradox Of Choice: Why More Is Less, January. https://works.swarthmore.edu/fac-psychology/198.

Tapal, Adam, Ela Oren, Reuven Dar, and Baruch Eitam. 2017. “The Sense of Agency Scale: A Measure of Consciously Perceived Control over One’s Mind, Body, and the Immediate Environment.” Frontiers in Psychology 8 (September). https://doi.org/10.3389/fpsyg.2017.01552.

Wu, Yuanjie, Yu Wang, Sungchul Jung, Simon Hoermann, and Robert W. Lindeman. 2019. “Exploring the Use of a Robust Depth-Sensor-Based Avatar Control System and Its Effects on Communication Behaviors.” In Proceedings

of the 25th ACM Symposium on Virtual Reality Software and Technology, 1–9. VRST '19. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3359996.3364267.

Zhi, Su, Peng Qunjie, Wang Chaolun, Xiang Wang, and Ling Chen. 2022. "The Influence of Figures in Warning Signs at the Manual Toll Station on the Lane Change Timing of Drivers in the Context of Virtual Reality of High-Proportion ETC Vehicles." Complexity 2022 (1): 7121952. https://doi.org/10.1155/2022/7121952.