A Formative Investigation of Student Perceptions of Digital Twin Instructors in Remote Lecture Environments

Matias Volonte
School of Computing, Synthetic Personas Lab
Clemson University
Charleston, South Carolina, USA
mvolont@clemson.edu

Nicholas Joseph Berns
Clemson University
Clemson, South Carolina, USA
nberns@g.clemson.edu

Carlos Toxtli
Human-AI Empowerment Lab
Clemson University
Clemson, South Carolina, USA
ctoxtli@clemson.edu

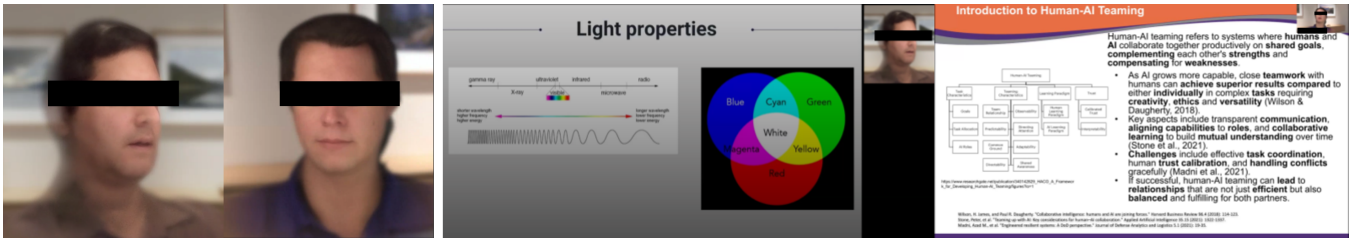

Figure 1: Comparison of a real instructor and their digital twin avatar delivering remote lectures via Zoom.

Abstract

The prevalence of remote instruction has heightened the need to understand how emerging avatar technologies might support instructor presence in virtual classrooms. Digital twins, defined here as animated avatars that mirror a speaker's facial expressions and movements in real time, represent a potential alternative to standard webcam video. This formative study examines student perceptions following exposure to a digital twin instructor during live Zoom lectures. Twenty-four students completed a post-session survey assessing perceived realism, emotional engagement, and willingness to attend future sessions. Spearman correlation analysis revealed that affective qualities such as likeability and enjoyment showed strong positive associations with receptivity, while perceptions of awkwardness and weirdness showed strong negative associations. Perceived natural movement was strongly associated with human resemblance and inversely associated with eeriness, consistent with uncanny valley theory. These exploratory findings identify candidate factors for future controlled investigations and offer preliminary design guidance emphasizing behavioral fidelity and emotional warmth over photorealism alone.

CCS Concepts

• **Human-centered computing** → **Virtual reality**; **Collaborative and social computing**; *User studies*; • **Applied computing** → *Interactive learning environments.*

Keywords

Digital Twin, Avatar, Remote Education, User Perception, Formative Study, Social Presence

ACM Reference Format:
Matias Volonte, Nicholas Joseph Berns, and Carlos Toxtli. 2026. A Formative Investigation of Student Perceptions of Digital Twin Instructors in Remote Lecture Environments. In *Extended Abstracts of the 2026 CHI Conference on Human Factors in Computing Systems (CHI EA '26), April 13–17, 2026, Barcelona, Spain.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3772363.3798975

CHI EA '26, Barcelona, Spain

ACM ISBN 979-8-4007-2281-3/26/04
https://doi.org/10.1145/3772363.3798975

1 Introduction

The widespread adoption of remote education has intensified scholarly attention to student engagement and instructor presence in virtual classrooms [15, 18]. Video conferencing platforms enable synchronous communication but impose constraints on nonverbal cues that contribute to social presence and learning outcomes [1, 4, 23]. Research on Zoom fatigue suggests prolonged video conferencing imposes cognitive costs that may undermine educational effectiveness [7, 22].

Digital twins represent an emerging approach to addressing these limitations (see Figure 1). We define digital twins as animated avatars that replicate a speaker's facial expressions and movements in real time, distinguishing them from pre-recorded virtual agents or stylized cartoon representations. Unlike conventional webcam video, digital twins can provide consistent visual presentation while preserving expressive communication. Avatar-based representations may also reduce self-consciousness associated with being on camera [2]. However, prior work demonstrates that deviations in realism can elicit discomfort associated with the uncanny valley phenomenon [17, 19], making user perceptions critical to understand.

Given the nascent state of this technology in educational contexts, formative research is needed to identify perceptual factors

that may influence student receptivity. This exploratory study investigates students' responses to a live lecture delivered through a digital twin avatar integrated with conventional video conferencing, addressing two research questions: **RQ1:** What perceptual factors are associated with students' willingness to attend future classes with a digital twin instructor? **RQ2:** What perceptual factors are associated with perceived natural movement of the avatar? These questions generate hypotheses for future controlled investigations rather than test causal relationships.

This work makes three contributions. First, it provides empirical observations from an authentic classroom deployment, bridging the gap between laboratory studies and practical implementation. Second, it identifies specific perceptual dimensions warranting attention in future research on avatar-mediated instruction. Third, it offers preliminary design guidance for researchers and practitioners exploring digital twin technologies in educational settings.

2 Related Work

2.1 Social Presence and Virtual Agents

Social presence theory posits that communication media vary in their capacity to convey the psychological sense that others are present [23]. Video conferencing constrains social and emotional dimensions of communication compared to face-to-face interaction [1, 20]. The concept of transformed social interaction [2] suggests avatar-mediated communication offers unique affordances, including consistent visual presentation across varying conditions.

Research on pedagogical agents demonstrates that animated characters can support learning when designed with attention to social cues [10, 14, 24]. The persona effect suggests lifelike characters can positively influence student engagement [16], and instructor immediacy behaviors remain important predictors of student outcomes in mediated contexts [25]. Systems such as ViGather [13] have explored photorealistic avatars in collaborative environments, finding that avatar quality affects perceived social presence. However, most existing research focuses on immersive VR or stylized avatars, leaving a gap in understanding digital twins integrated with conventional video conferencing tools.

2.2 The Uncanny Valley

The uncanny valley phenomenon describes how near-human representations can elicit discomfort when subtle imperfections break the illusion of humanity [19]. Research establishes that motion quality is particularly critical in triggering uncanny responses [12, 17, 21]. Ho and MacDorman developed indices for measuring perceived humanness and eeriness [11]. Studies find that behavioral realism significantly affects willingness to engage with virtual agents [3, 8]. This study addresses the gap by examining real-time digital twins deployed within conventional video conferencing for classroom instruction.

3 Methodology

3.1 Study Design

This formative study examined student perceptions following exposure to a digital twin instructor during live remote lectures. Consistent with established approaches to formative HCI research [9, 26], we employed an exploratory design without a control condition or preregistered hypotheses. All reported associations are correlational and should be interpreted as identifying patterns worthy of future investigation rather than establishing causal relationships. The absence of a control condition means findings speak to variation in perceptions within the digital twin experience rather than comparing digital twins to conventional video.

3.2 Technical Implementation

The digital twin system was controlled using Apple Vision Pro, which captured the instructor's facial expressions and head movements via built-in sensors. Custom software mapped detected facial action units to a high-fidelity 3D avatar rendered in real time. The avatar video feed was shared as the instructor's camera feed within Zoom. The system operated at approximately 30 fps with total latency of 100 to 150 milliseconds. The avatar tracked major facial regions and head rotation, though fine-grained expressions such as micro-expressions were not fully captured.

3.3 Procedure and Participants

Two professors delivered lectures to their respective classes while represented by digital twin avatars via Zoom, with lectures lasting approximately 50 minutes. Students were not informed in advance that the instructor would appear as a digital twin, preserving ecological validity. Following each session, students completed a voluntary, anonymous survey. The study was approved by the institutional review board. A total of 24 students completed the survey, consistent with formative HCI studies [5].

3.4 Measures

The survey employed single-item 7-point Likert scales assessing constructs identified in prior work [11, 19]. Items were organized into four categories: Perceived Realism (natural movement, facial realism, human-likeness, smoothness, natural behavior, felt real), Negative Perceptual Qualities (robotic, awkward, unnatural, eerie, weird), Emotional Engagement (comfortable, pleasant, comfortable interacting, warm, likeable, enjoyable, engaging), and Repetition Desirability (willingness to interact again, willingness to attend future classes). Single-item measures were selected to minimize participant burden; future confirmatory research should employ validated multi-item instruments [6]. Table 3 provides descriptive statistics.

3.5 Data Analysis

Given the small sample and non-normal distributions, Spearman rank correlations were computed. Only correlations achieving $p < .05$ are reported. Consistent with formative research, no multiple comparison corrections were applied; results should be interpreted as hypothesis-generating.

4 Results

Descriptive statistics appear in Table 3. Students rated the avatar moderately on perceived realism (Natural Movement: M = 4.08, SD = 1.74) and engagement (Likeability: M = 5.08, SD = 1.29). Negative qualities such as Awkwardness (M = 4.21) and Weirdness (M = 4.58) were rated near or above the midpoint. Notably, Repetition

Table 1: Spearman correlations with Repetition Desirability (p < .05).

Variable	ρ	p
Positive		
Likeability	.768	<.001
Comfortable Interaction	.766	<.001
Enjoyment	.765	<.001
Interact Again	.714	<.001
Pleasantness	.639	.001
Facial Realism	.606	.002
Smoothness of Motion	.558	.005
Natural Behavior	.532	.007
Warmness	.413	.045
Negative		
Dislikeability	-.770	<.001
Weirdness	-.629	.001
Coldness	-.619	<.001
Eeriness	-.520	.009
Awkwardness	-.477	.018
Robotic Impression	-.456	.005

Table 2: Spearman correlations with Perceived Natural Movement (p < .05).

Variable	ρ	p
Positive		
Human Resemblance	.882	<.001
Natural Behavior	.817	<.001
Comfortableness	.779	<.001
Felt Real	.775	<.001
Facial Realism	.771	<.001
Smoothness of Motion	.768	<.001
Pleasantness	.757	<.001
Likeability	.694	<.001
Repetition Desirability	.521	.011
Enjoyment	.464	.026
Negative		
Eeriness	-.844	<.001
Awkwardness	-.792	<.001
Unnaturalness	-.784	<.001
Weirdness	-.749	<.001
Dislikeability	-.728	<.001
Robotic Impression	-.686	<.001

Desirability was relatively low (M = 3.00, SD = 1.29), indicating limited enthusiasm for repeating the experience.

4.1 RQ1: Factors Associated with Repetition Desirability

Table 1 presents correlations with Repetition Desirability. Willingness to attend future sessions showed strong positive associations with Likeability (ρ = .768), Enjoyment (ρ = .765), and Comfortable Interaction (ρ = .766). Perceptual realism variables also showed positive associations, including Facial Realism (ρ = .606) and Natural Behavior (ρ = .532). Repetition Desirability showed strong negative associations with Dislikeability (ρ = -.770), Weirdness (ρ = -.629), Coldness (ρ = -.619), and Awkwardness (ρ = -.477). These patterns suggest receptivity depends on both achieving positive affective impressions and minimizing uncanny perceptions.

4.2 RQ2: Factors Associated with Natural Movement

Table 2 presents correlations with Perceived Natural Movement. Natural Movement was strongly associated with Human Resemblance (ρ = .882), Natural Behavior (ρ = .817), and Comfortableness (ρ = .779). Strong negative associations with Eeriness (ρ = -.844), Awkwardness (ρ = -.792), and Unnaturalness (ρ = -.784) are consistent with uncanny valley theory [19]. Natural Movement was also significantly associated with Repetition Desirability (ρ = .521), suggesting motion quality may be important in overall receptivity.

5 Discussion

5.1 Interpretation of Findings

The correlation patterns suggest that emotional response may be more predictive of acceptance than technical realism alone. The strong associations between affective variables (likeability, enjoyment) and willingness to attend future sessions align with research on pedagogical agents [16, 24]. The strong negative associations between uncanny perceptions (eeriness, awkwardness) and both natural movement and repetition desirability are consistent with uncanny valley theory [19]. The particularly strong correlation between natural movement and eeriness (ρ = -.844) supports findings that motion quality is critical in avatar acceptance [12].

The relatively low Repetition Desirability (M = 3.00) indicates current technology may not yet achieve the quality threshold needed for widespread acceptance. However, substantial variance and strong correlations suggest some students responded positively, and improving identified factors could increase acceptance.

5.2 Design Implications

These formative findings suggest several preliminary design considerations. First, developers should prioritize behavioral fidelity and motion quality over static visual realism. Second, avatar design should attend to warmth and emotional expressiveness. Third, designers should actively minimize uncanny valley effects through testing with target populations.

5.3 Limitations

Several limitations bound interpretation. The small sample (N = 24) limits generalizability. The absence of a control condition prevents comparison with webcam video. The correlational design precludes causal inference. Single-item measures introduce unknown measurement error. Future research should address these limitations through controlled experiments, validated multi-item measures, and objective learning outcome measures.

6 Conclusion

This formative study examined student perceptions of a digital twin instructor during live Zoom lectures. Affective qualities such as likeability and enjoyment showed strong positive associations with

Table 3: Descriptive statistics (N = 24). All items used 7-point scales.

Variable	M	SD
Perceived Realism		
Natural Movement	4.08	1.74
Natural Behavior	4.29	1.57
Felt Real	3.58	1.93
Facial Realism	2.88	1.76
Smoothness of Motion	4.33	1.66
Human Resemblance	4.50	1.78
Negative Qualities		
Robotic Impression	4.04	1.52
Awkwardness	4.21	1.69
Unnaturalness	4.08	1.64
Eeriness	3.04	1.80
Weirdness	4.58	2.02
Emotional Engagement		
Comfortableness	4.54	1.70
Pleasantness	4.08	1.79
Comfortable Interaction	4.50	1.70
Warmness	3.46	1.38
Likeability	5.08	1.29
Enjoyment	4.25	1.73
Engagement	4.58	1.57
Repetition		
Interact Again	4.29	1.78
Repetition Desirability	3.00	1.29

receptivity, while uncanny perceptions including awkwardness and weirdness showed strong negative associations. These patterns suggest avatar design for educational contexts should prioritize behavioral fidelity and emotional warmth alongside technical realism. The exploratory nature of this research means findings should be treated as hypotheses warranting future controlled investigation. Nevertheless, the real-world deployment context provides ecological validity, and identified patterns offer preliminary guidance for researchers and practitioners exploring avatar-mediated instruction.

References

[1] Jeremy N Bailenson. 2021. Nonverbal overload: A theoretical argument for the causes of Zoom fatigue. *Technology, Mind, and Behavior* 2, 1 (2021). doi:10.1037/tmb0000030

[2] Jeremy N Bailenson, Andrew C Beall, Jack Loomis, Jim Blascovich, and Matthew Turk. 2004. Transformed social interaction: Decoupling representation from behavior and form in collaborative virtual environments. *Presence: Teleoperators & Virtual Environments* 13, 4 (2004), 428–441.

[3] Jeremy N Bailenson, Nick Yee, Dan Merget, and Ralph Schroeder. 2006. The effect of behavioral realism and form realism of real-time avatar faces on verbal disclosure, nonverbal disclosure, emotion recognition, and copresence in dyadic interaction. *Presence: Teleoperators and Virtual Environments* 15, 4 (2006), 359–372.

[4] Frank Biocca, Chad Harms, and Judee K Burgoon. 2003. Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence: Teleoperators & Virtual Environments* 12, 5 (2003), 456–480. doi:10.1162/105474603322761270

[5] Kelly Caine. 2016. Local Standards for Sample Size at CHI. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. 981–992. doi:10.1145/2858036.2858498

[6] Adamantios Diamantopoulos, Marko Sarstedt, Christoph Fuchs, Petra Wilczynski, and Sebastian Kaiser. 2012. Guidelines for Choosing Between Multi-Item and Single-Item Scales for Construct Measurement: A Predictive Validity Perspective. *Journal of the Academy of Marketing Science* 40, 3 (2012), 434–449. doi:10.1007/s11747-011-0300-3

[7] Géraldine Fauville, Mufan Luo, Anna C. M. Queiroz, Jeremy N. Bailenson, and Jeff Hancock. 2021. Zoom Exhaustion & Fatigue Scale. *Computers in Human Behavior Reports* 4 (2021), 100119. doi:10.1016/j.chbr.2021.100119

[8] Maia Garau, Mel Slater, Vinoba Vinayagamoorthy, Andrea Brogni, Anthony Steed, and M Angela Sasse. 2003. The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 529–536.

[9] William Gaver. 2012. What Should We Expect from Research Through Design? (2012), 937–946. doi:10.1145/2207676.2208538

[10] Arthur C Graesser, Shulan Lu, George Tanner Jackson, Heather Hite Mitchell, Mathew Ventura, Andrew Olney, and Max M Louwerse. 2004. AutoTutor: A tutor with dialogue in natural language. *Behavior Research Methods, Instruments, & Computers* 36, 2 (2004), 180–192.

[11] Chin-Chang Ho and Karl F MacDorman. 2017. Revisiting the uncanny valley theory: Developing and validating an alternative to the Godspeed indices. *Computers in Human Behavior* 26, 6 (2017), 1508–1518.

[12] Jessica K Hodgins, Wayne L Wooten, David C Brogan, and James F O'Brien. 1998. Animating human athletics. In *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques*. ACM, 71–78.

[13] Dongsik Jo, Kangsoo Kim, Greg F Welch, Myounghoon Jeon, and Thomas D Parsons. 2023. ViGather: Inclusive Virtual Conferencing with a Joint Experience of 2D and VR Users. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. ACM, 1–16. doi:10.1145/3544548.3581218

[14] W. Lewis Johnson, Jeff W. Rickel, and James C. Lester. 2000. Animated Pedagogical Agents: Face-to-Face Interaction in Interactive Learning Environments. *International Journal of Artificial Intelligence in Education* 11, 1 (2000), 47–78.

[15] Kyungbin Kim, Kyungwon Kwon, Anne Ottenbreit-Leftwich, Krista Glazewski, and Thomas Brush. 2021. Learning design to support student-AI collaboration: Perspectives of leading teachers for AI in education. *Education and Information Technologies* 26, 4 (2021), 4921–4941.

[16] James C Lester, Sharolyn A Converse, Susan E Kahler, S Todd Barlow, Brian A Stone, and Ravinder S Bhogal. 1997. The persona effect: Affective impact of animated pedagogical agents. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems*. ACM, 359–366.

[17] Karl F MacDorman and Hiroshi Ishiguro. 2006. Subjective ratings of robot video clips for human likeness, familiarity, and eeriness: An exploration of the uncanny valley. *ICCS/CogSci-2006 Long Symposium: Toward Social Mechanisms of Android Science* (2006), 26–29.

[18] Florence Martin and Doris U. Bolliger. 2018. Engagement Matters: Student Perceptions on the Importance of Engagement Strategies in the Online Learning Environment. *Online Learning* 22, 1 (2018), 205–222. doi:10.24059/olj.v22i1.1092

[19] Masahiro Mori, Karl F MacDorman, and Norri Kageki. 2012. The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine* 19, 2 (2012), 98–100. doi:10.1109/MRA.2012.2192811

[20] Catherine S. Oh, Jeremy N. Bailenson, and Gregory F. Welch. 2018. A Systematic Review of Social Presence: Definition, Antecedents, and Implications. *Frontiers in Robotics and AI* 5 (2018), 114. doi:10.3389/frobt.2018.00114

[21] Lukasz Piwek, Lawrie S. McKay, and Frank E. Pollick. 2014. Empirical Evaluation of the Uncanny Valley Hypothesis Fails to Confirm the Predicted Effect of Motion. *Cognition* 130, 3 (2014), 271–277. doi:10.1016/j.cognition.2013.11.001

[22] René Riedl. 2022. On the stress potential of videoconferencing: Definition and root causes of Zoom fatigue. *Electronic Markets* 32, 1 (2022), 153–177.

[23] John Short, Ederyn Williams, and Bruce Christie. 1976. *The Social Psychology of Telecommunications*. John Wiley & Sons, London.

[24] George Veletsianos, Robert Heller, Scott Overmyer, and Mike Procter. 2010. Conversational agents in virtual worlds: Bridging disciplines. *British Journal of Educational Technology* 41, 1 (2010), 123–140.

[25] Paul L Witt, Lawrence R Wheeless, and Mike Allen. 2004. A meta-analytical review of the relationship between teacher immediacy and student learning. *Communication Monographs* 71, 2 (2004), 184–207.

[26] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. 2007. Research Through Design as a Method for Interaction Design Research in HCI. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 493–502. doi:10.1145/1240624.1240704