

SPECIAL ISSUE PAPER OPEN ACCESS

CASAXR 26: International Conference on Computer Animation, Social Agents, and Extended Reality

Effects of Immersion Fidelity and Multimodal Support on Cognitive Load, User Experience, and Behavioral Intentions in Embodied Conversational Agent Interactions

Sagar A. Vankit¹  | Nafisa Sabnam Shama¹  | Jay Ashish Kadam¹  | Vivian Genaro Motti²  | Joseph LaViola³  | Andrew Duchowski¹  | Matias Volonte¹ 

¹School of Computing, Clemson University, Clemson, SC, USA | ²Department of Information Sciences and Technology, George Mason University, Fairfax, VA, USA | ³Department of Computer Science, University of Central Florida, Orlando, FL, USA

Correspondence: Sagar A. Vankit (svankit@clemson.edu)

Received: 16 April 2026 | **Revised:** 12 May 2026 | **Accepted:** 14 May 2026

Keywords: cognitive load | embodied conversational agents | health counseling | virtual reality | working alliance

ABSTRACT

This study examines how immersion fidelity (2D vs. VR) and multimodal support shape user experience during 20+ min diabetes-related conversations with an embodied conversational agent (ECA). Using a between-subjects design, participants engaged in conversation-only interactions or conversations augmented with subtitles, visualizations, or both. We assessed memory retention, diet attitude change, behavioral commitment intentions to adopt healthier eating habits, perceived workload, working alliance, and subjective impressions of the agent. In conversation-only interactions, 2D supported stronger working alliance, while VR elicited stronger behavioral commitment intentions. When subtitles and/or visualizations were present, working alliance and behavioral commitment intentions converged across modalities, although subjective impressions remained condition-specific. Across conditions, VR was associated with higher perceived workload, including consistently higher physical demand and additional increases in mental and overall workload under multimodal support. Memory retention and diet attitude change did not differ between 2D and VR. These findings highlight tradeoffs between relational alignment, motivational impact, and interaction effort when designing scalable ECA-mediated health conversations across 2D and VR platforms.

1 | Introduction

Face-to-face health counseling is widely regarded as the gold standard for preventive care, chronic disease management, and health behavior change. In-person interactions enable providers to build rapport, deliver personalized feedback, and sustain patient motivation [1]. Large-scale clinical trials, including the Diabetes Prevention Program, show that structured lifestyle counseling and dietary support can significantly reduce the incidence of conditions such as type 2 diabetes, hypertension, and obesity [2]. However, in-person counseling is resource-intensive, requiring trained professionals, frequent

sessions, and substantial healthcare infrastructure [3], which limits scalability and access. This has created a need for technology-mediated alternatives that can extend the benefits of traditional counseling while expanding reach and maintaining engagement [4].

Embodied conversational agents (ECAs) offer a promising approach to scalable health communication. By combining verbal and nonverbal behaviors, ECAs simulate face-to-face dialogue and have been shown to foster trust, empathy, and relational continuity [5, 6]. Using motivational interviewing strategies, ECAs can motivate health behavior change across repeated

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2026 The Author(s). *Computer Animation and Virtual Worlds* published by John Wiley & Sons Ltd.

interactions [7, 8], while remaining cost-effective and broadly deployable, including in settings with limited access to in-person care [9]. These properties make ECAs attractive for extending evidence-based practices in preventive health and lifestyle coaching [10, 11].

However, sustained, long-term interaction has long been recognized as important in health-oriented relational agent research, particularly for supporting relationship building, reflection, and engagement over time [12]. Counseling and motivational interviewing research likewise emphasize that effective interventions unfold over time, progressing through stages of rapport-building, reflection, and goal alignment to support health behavior change [13–16]. Supporting longer, more natural ECA conversations is therefore critical for aligning digital interventions with established counseling practices.

In addition, ECA-based health interactions have been studied in both 2D and VR interfaces [17, 18], yet few studies directly compare these modalities systematically within matched, extended counseling sessions lasting 20 min or more. Moreover, existing work rarely examines how key outcome dimensions such as cognitive workload, relational alignment, memory, and behavioral intentions may diverge rather than move in parallel as immersive fidelity increases. This study addresses these gaps by directly comparing extended ECA-based health counseling delivered in 2D and VR environments to examine how immersive fidelity and multimodal support differentially shape cognitive, relational, and behavior-relevant outcomes during sustained conversational interactions.

1.1 | Study Contributions

- **Extended cross-modality comparison with multimodal manipulation.** This study provides a controlled comparison of embodied health counseling delivered in 2D and VR during sustained (20+ min) conversational interactions, while systematically examining how subtitles and visualizations, individually and in combination, interact with immersive fidelity to shape user experience.
- **Condition-dependent modality effects in sustained ECA interactions.** Modality-related differences in working alliance, behavioral commitment intentions to adopt healthier eating habits, and subjective impressions emerged primarily in minimally supported, conversation-only interactions. Working alliance and behavioral commitment differences were attenuated or eliminated when subtitles, visualizations, or their combination were present, indicating that multimodal support moderates modality effects rather than producing uniform platform differences.
- **Distinct workload tradeoffs across immersive fidelities.** Immersive VR was associated with higher perceived physical workload across conditions, with additional increases in mental and overall workload emerging under multimodal support. These findings highlight tradeoffs between immersive interaction demands and user effort rather than a single optimal platform.

- **Immersive VR supports comparable health outcomes through distinct experiential pathways.** While core health-related outcomes such as memory retention and diet attitude change were comparable across 2D and VR platforms, immersive VR achieved these outcomes alongside stronger behavioral commitment intentions to adopt healthier eating habits in minimally supported interactions and distinct perceptual response patterns, demonstrating that immersive fidelity can motivate health intentions through different experiential pathways.

2 | Related Work

2.1 | Multimedia Learning and Support Features

Research on multimedia learning provides a framework for understanding how subtitles and visualizations influence comprehension and recall. The cognitive theory of multimedia learning posits that people process verbal and visual information through separate channels with limited capacity, and that learning improves when information is integrated across both [19]. Evidence shows that same-language captions can enhance comprehension and memory by reinforcing auditory information [20], while visualizations such as diagrams or annotated images clarify abstract concepts and direct attention to critical details [21, 22]. Likewise, poorly integrated materials can produce split-attention effects that increase cognitive load and hinder retention [23, 24].

2.2 | Memory Outcomes in VR and 2D Environments

Research comparing immersive VR with desktop settings has shown that while VR can increase presence, some studies report lower factual learning under certain task conditions [25]. Other work similarly finds that VR learners may recall less content and transfer knowledge less effectively than those in 2D conditions, despite higher levels of engagement [26]. By contrast, evidence from 2D and other non-immersive settings more consistently shows benefits when multimedia supports are introduced, suggesting that learning outcomes depend strongly on task structure and design choices.

2.3 | Cognitive Load in 2D and VR Platforms

Prior work consistently shows that 2D platforms are associated with lower perceived workload than immersive VR. Desktop conditions have been linked to reduced effort and more efficient recall relative to VR [26], and neurophysiological evidence indicates lower neural markers of cognitive effort when tasks are performed on 2D displays [27]. Comparisons of display types further suggest that while VR enhances motivation and embodiment, it does so at the cost of increased processing demands [28].

More broadly, immersive VR reliably increases attentional engagement but is also associated with higher cognitive and

physical demands than less immersive media [29]. Elevated presence requires users to allocate resources to richer sensory input, and when multiple elements compete for attention, workload can increase further, contributing to visual fatigue and discomfort [30]. These findings underscore a recurring trade-off between engagement and processing effort across immersive platforms.

2.4 | Interaction Differences, Perceptual Responses, and Relational Outcomes Across 2D and VR

When interacting with embodied conversational agents, users' experiences can differ between 2D and VR. In 2D settings, interfaces often support smoother turn-taking and lower-effort interaction, as users face fewer attentional demands and can process the agent's cues more efficiently [27, 31]. By contrast, VR presentations heighten the sense of co-location and realism, frequently increasing engagement while also drawing on additional cognitive resources [18, 32].

Social presence. In VR, users typically report stronger feelings of social presence and interpersonal salience [29], which can amplify engagement and influence dialogue behaviors such as gaze alignment [32]. While this heightened presence can enhance affective involvement, it does not uniformly translate into superior learning or performance outcomes across tasks [26].

Affective impressions. Users in 2D often perceive agents as more predictable and less demanding, which can support relational stability. VR, in contrast, may evoke stronger affective reactions. Increased realism and immersion enhance emotional arousal, but they can also make users more sensitive to behavioral mismatches, which may influence trust or rapport [33].

Perceptual responses. In 2D environments, agents are processed in a less demanding perceptual frame, reducing the risk of overload. VR, however, can intensify perceptual mental and physical strain when multiple visual elements compete for processing [27, 30]. These patterns reflect a trade-off: while VR magnifies presence and affective intensity, 2D often affords more cognitively efficient and stable interactions.

Working alliance. Relational outcomes also vary by modality. In 2D, lower perceptual demands and more stable cues can enable stronger emotional bonds, smoother task agreement, and clearer goal alignment [12]. VR, by contrast, raises expectations for behavioral realism; when gaze and gesture cues are well-timed, alliance can be amplified, while mismatches in timing or nonverbal fidelity may attenuate rapport [33, 34]. Meta-analytic evidence from psychotherapy underscores that alliance is a strong predictor of outcomes generally [35], making these modality-driven differences especially important for the design of health ECAs.

3 | System Design and Architecture

Although the 2D condition used the same three-dimensional ECA model, it was presented through a screen-based display rather than an immersive environment; therefore, the comparison concerns platform immersive fidelity rather than the

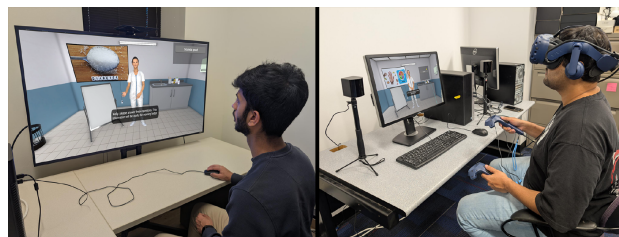


FIGURE 1 | Participant in the 2D condition (left) and VR condition (right).

dimensionality of the character model. The study was conducted in controlled laboratory settings using both a 2D screen-based environment and an immersive VR environment. Sessions took place in quiet rooms to minimize external distractions. In the 2D modality, the system was displayed on a 55-in television connected to a Windows desktop computer. The 55-in display was selected to approximate a more life-sized interaction with the virtual agent, consistent with prior peer-reviewed virtual human studies that used large-screen setups to support more realistic interpersonal encounters [36]. In the VR modality, participants used an HTC Vive Pro Eye head-mounted display and interacted within a three-dimensional consultation room. Across both modalities, the system ran on identical hardware, including an Intel Core i7-9700 processor, an NVIDIA GeForce RTX 2080 Ti GPU, 64 GB of RAM, a standard keyboard, and external speakers (Figure 1).

3.1 | System Description

The system was designed to support controlled yet naturalistic dialogue with an embodied conversational agent across both modalities. The conversational scripts were developed and reviewed in consultation with a qualified medical professional to ensure clinical accuracy, appropriateness, and relevance of the health-related information. Interaction followed a structured dialogue flow to ensure consistency in content, timing, and conversational logic. After each agent utterance, participants selected from a menu of predefined response options that allowed them to advance the conversation or request repetition (See Figure 2). This constrained interaction design ensured safety and accuracy in the delivery of health-related information by preventing unrestricted free-text input as noted in previous studies [37].

3.2 | 2D and VR Setups

The primary distinction between conditions was the mode of presentation. In the 2D condition, participants interacted with the agent on a large desktop display, whereas in the VR condition they were immersed in a three-dimensional consultation room rendered through a head-mounted display. The VR environment was optimized for stable rendering above 90 frames per second to support smooth interaction. In both modalities, identical conversational content, response options, and interface elements were presented, with the conversational menu displayed either on the screen or within the virtual space (Figures 1 and 2). Aside from differences in immersive fidelity, all aspects of the interaction were held constant across conditions.

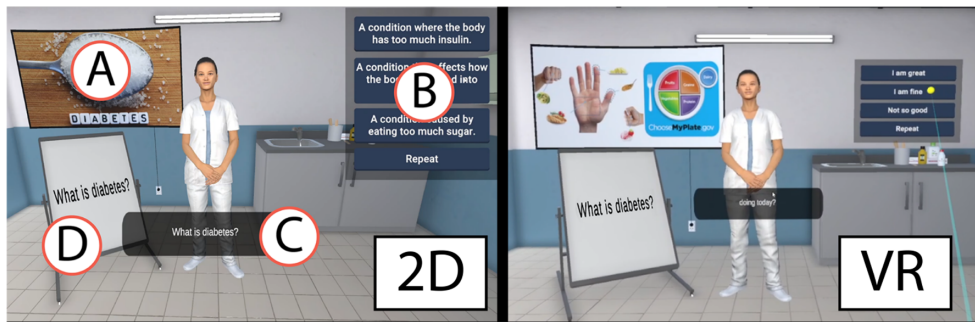


FIGURE 2 | Interface components available across conditions: (A) Image visualization, (B) Conversational menu, (C) Subtitles panel, and (D) Text visualization, presented in 2D and VR.

3.3 | ECA's Verbal and Non-Verbal Behaviors

The agent, named Sarah, functioned as a healthcare coach focusing on diabetes prevention and management through improved dietary habits. Conversations were structured around motivational interviewing techniques, emphasizing empathy, autonomy, and reflective engagement to encourage intentions to adopt healthier behaviors [16].

To reinforce communication, Sarah displayed both verbal and non-verbal behaviors. A Microsoft Rocketbox character's facial expressions were modeled using blendshapes based on the Facial Action Coding System [38, 39], while speech was synchronized through the Unity animation tool. Non-verbal behaviors included head nods, pointing to visual aids, open-palm gestures, and posture shifts, all aligned with conversational content. These expressive features, combined with motivational interviewing strategies, were reviewed by a medical specialist to ensure clarity, realism, and accuracy of the health information presented.

4 | Research Questions and Hypotheses

We developed the following research questions and corresponding hypotheses:

4.1 | RQ1

Do conversations in VR with the ECA lead to higher perceived workload compared to conversations in a 2D platform?

4.2 | H1

Participants engaging in conversations in the VR environment will report higher perceived workload than those in the 2D interface on NASA TLX measures, with modality effects varying across workload dimensions and multimodal support conditions.

Rationale: VR environments typically impose greater cognitive and physical demands to users due to increased sensory input, spatial processing, and embodied interaction. These demands can elevate perceived workload even when task content remains constant. Prior work has shown that immersive VR can increase cognitive load compared to desktop or 2D presentations [25].

4.3 | RQ2

How do conversations on a 2D platform versus in a VR environment affect the formation of a working alliance with an embodied conversational agent?

4.4 | H2

Participants in the 2D condition will report stronger working alliance, particularly in emotional bond and perceived clarity of interaction, compared to those in the VR condition.

Rationale: Gratch et al. [34], argues that in immersive VR, users hold higher expectations for behavioral realism, making them more sensitive to mismatches in timing, motion, or non-verbal expression. These mismatches can reduce liking and trust and may weaken rapport-related responses in immersive interactions [40]. In contrast, 2D agents reduce the risk of such perceptual disruption and may offer a more stable interaction frame, enabling stronger feelings of trust, emotional connection, and shared task focus.

4.5 | RQ3

How do conversations on a VR platform versus a 2D platform influence participants' behavioral commitment intentions to adopt healthier eating habits following an ECA-led health conversation?

4.6 | H3

Participants in the VR condition will report stronger behavioral commitment intentions to adopt healthier eating habits compared to those in the 2D condition.

Rationale: Immersive VR environments have been shown to enhance motivational processes by increasing affective arousal and experiential realism [41].

4.7 | RQ4

Does the use of a VR platform influence memory retention of health-related information compared to a 2D interface?

4.8 | H4

Memory retention following an ECA-led health conversation may differ between VR and 2D platforms depending on the presence and type of multimodal support.

Rationale: Prior studies suggest that VR can enhance cognitive engagement and memory encoding through heightened immersion, spatial presence, and attentional focus [25]. The embodied and interactive nature of VR is theorized to support deeper information processing, potentially leading to improved memory recall.

5 | Experiment Methodology

5.1 | Study Design

We employed a 2 (Mode: 2D vs. VR) $\times$ 4 (Multimodal Condition: Conversation Only (CON), Conversation with Subtitles (CSU), Conversation with Visualizations (CVI), and Conversation with both Subtitles and Visualizations (CSV)) between-subjects experimental design (See Figure 3 and Table 1.) Participants were randomly assigned to one of the eight conditions.

5.2 | Conversation Topic

The conversational agent engaged participants in a structured discussion about diabetes prevention and healthy dietary practices. The dialogue covered topics such as daily eating habits, portion control, and strategies for reducing added sugars and unhealthy fats. The content was adapted from dietary guidelines provided by *MyPlate.gov* and was reviewed by a licensed medical doctor specializing in metabolic health [42]. The conversation was designed to last more than 20 min, providing sufficient time for reflection, self-assessment, and personalized goal-setting, consistent with motivational interviewing frameworks in which behavior change emerges through sustained dialogue rather than brief informational exchanges [15, 43].

5.3 | Experimental Conditions

Participants were assigned to one of four multimodal support conditions and interacted with the ECA in either a 2D or VR modality. Across all conditions and modalities, conversational content, agent behavior, timing, and interaction structure were held constant; only the presence of subtitles and visualizations differed (See Figure 3 and Table 1).

5.4 | Participants

A total of 160 participants were recruited from a university campus and randomly assigned to one of eight groups (2D or VR $\times$ CON, CSU, CVI, or CSV; $N = 20$ per group). Each modality condition (2D or VR) included 80 participants.

In the 2D condition, participants had a mean age of 24.71 years ($SD = 3.70$), and the sample included 55 men and 25 women. The racial composition was 31.25% White (not Hispanic), 35% Asian or Pacific Islander, 8.75% Black (not Hispanic), 6.25% Hispanic or Latino, and 18.75% identifying as Other or Multiracial.

In the VR condition, participants had a mean age of 29.25 years ($SD = 6.81$), and the sample included 46 men and 34 women. The racial composition was 50.0% White (not Hispanic), 38.75% Asian or Pacific Islander, 7.5% identifying as Other, and 3.75% Black (not Hispanic).

All participants received a \$10 gift card for their participation.

5.5 | Power Analysis

A sensitivity analysis was conducted using G*Power v3.1 for two-group between-modality comparisons (2D vs. VR). For comparisons within each multimodal support condition ($n = 20$ per group), the design provided 80% power ($\alpha = .05$, two-tailed) to detect effects of approximately $d = 0.90$ or larger, corresponding to moderate-to-large differences. When analyses were collapsed across multimodal conditions ($n = 80$ per modality), the design provided 80% power to detect effects of approximately $d = 0.44$, corresponding to small-to-moderate differences. Effect sizes are therefore reported alongside p -values to support interpretation of the magnitude and practical relevance of observed effects.

5.6 | Study Procedure

Participants first listened to a brief overview of the study, provided informed consent, and confirmed their English fluency before

TABLE 1 | Experimental conditions.

Condition	Code	Sub	Viz
Conversation only	CON	No	No
Conversation + subtitles	CSU	Yes	No
Conversation + visualizations	CVI	No	Yes
Conversation + subtitles + visualizations	CSV	Yes	Yes

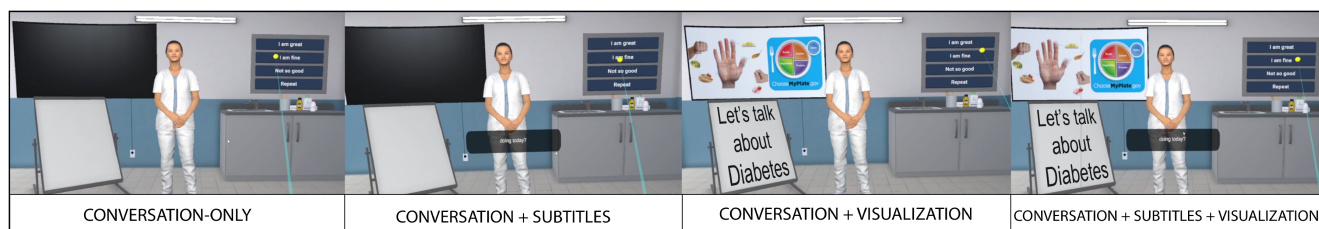


FIGURE 3 | Experimental conditions across 2D and VR platforms.

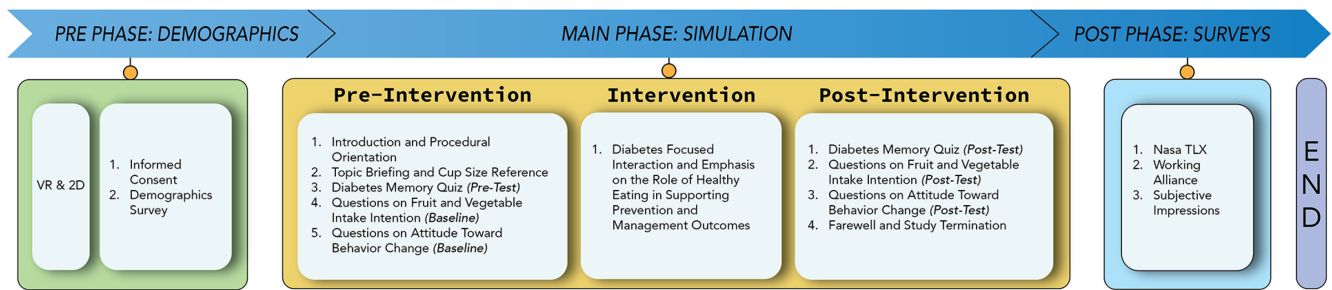


FIGURE 4 | Phases of the experiment, including the measures collected in each phase.

beginning. The study proceeded in three stages: a *Pre-Experiment Phase*, a *Main Experiment Phase*, and a *Post-Experiment Phase* (See Figure 4).

Pre-experiment Phase: Participants first completed a demographic questionnaire. Those in the VR condition then wore the HTC Vive Pro Eye headset to begin the VR portion of the experiment, while those in the 2D condition used a 55-inch TV display that served as the interface for the main experiment.

Main Experiment Phase: The ECA, introduced as Sarah, explained the study purpose and discussed diabetes and healthy eating. Participants then completed a 10-item diabetes quiz and a dietary attitudes survey assessing fruit consumption and intentions to adopt healthier eating habits, and how motivated they were to change their eating habits. This was followed by an interactive conversation with the ECA focused on diet and disease prevention. After the interaction, the same knowledge and attitude measures were administered again to evaluate retention and changes in behavioral intentions.

Post-Experiment Phase: Participants completed the NASA Task Load Index, and provided their subjective impressions.

5.7 | Measures

5.7.1 | Memory Outcome

Participants completed a 10-item diabetes knowledge quiz before and after the interaction with the ECA. Pre-post change scores were used for analysis.

5.7.2 | Diet Attitude

Participants reported daily fruit and vegetable intake by selecting from “None,” “1/2 cup,” “1 cup,” “1.5 cups,” or “2 cups or more,” which were converted to numerical values (0, 0.5, 1, 1.5, and 2, respectively). This cup-based format was chosen because portion sizes in cups were explicitly explained during the conversation, making the response scale directly consistent with the educational content delivered by the agent. The format also reflects a simple and commonly used dietary reporting approach aligned with public health nutrition guidance, allowing participants to estimate intake in a practical and interpretable way within a short experimental session. This measure was collected before and after

the interaction, and pre-post change scores indexed diet attitude change.

5.7.3 | Behavioral Commitment Intentions to Adopt Healthier Eating Habits

Participants rated their behavioral commitment intentions on a 7-point Likert scale before and after the interaction. Pre-post change scores were used for analysis. Behavioral commitment intentions are widely used as proximal indicators of health behavior change, as attitudes and intentions have been shown to be positively correlated with subsequent behavior across health domains [44, 45].

5.7.4 | NASA-TLX

Perceived workload was assessed using the NASA Task Load Index (NASA-TLX), which measures workload across six dimensions: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration. Each subscale was rated on a 0–100 scale. Following standard NASA-TLX procedures, participants completed pairwise comparisons to derive individual weights for each dimension. A weighted total workload score was then computed as a normalized weighted average of the six subscales. As a result, total workload values are reported on a reduced composite scale rather than the raw 0–100 TLX range, and comparisons focus on relative differences between conditions rather than absolute workload magnitude.

5.7.5 | Working Alliance

The Working Alliance Inventory measures the perceived quality of the collaborative relationship between a user and an interaction partner. Is divided in: *Task* (perceived agreement on goals and relevance of the activity), *Goal* (shared understanding of the intended outcome), and *Bond* (sense of mutual trust and emotional connection) [46].

5.7.6 | Subjective Impressions

Participants rated their agreement with a set of statements about the agent using a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree), with higher values

indicating stronger endorsement of each statement. The items were designed to capture perceptual, affective, and behavioral impressions of the agent. Items included: *Sarah was lively* (LIVELY), *I was comfortable talking to Sarah* (COMFORT), *Sarah was appropriate for the job* (APPROP), *Sarah was engaging* (ENGAGE), *Sarah was awkward* (AWKWARD), *I liked talking to Sarah* (LIKENESS), *I disliked talking to Sarah* (DISLIKENESS), *It was fun talking to Sarah* (FUN), *Sarah was eerie* (EERIE), *Sarah was weird* (WEIRD), *Sarah was cold* (COLD), *Sarah was unnatural* (UNNAT), *Sarah was energizing* (ENERGIZED), *I paid attention to Sarah* (ATTENTION), and *I ignored Sarah* (IGNORED).

6 | Results

6.1 | Comparative Analysis between Modalities Across Conditions

We compared user responses between the 2D and VR modalities using nonparametric tests using R. Analyses were conducted separately for each multimodal support condition, yielding four matched comparisons: 2D CON vs. VR CON, 2D CSU vs. VR CSU, 2D CVI vs. VR CVI, and 2D CSV vs. VR CSV. Memory outcomes, diet intake attitudes, behavioral commitment intentions to adopt healthier eating habits, working alliance, and subjective impression measures were analyzed using Kruskal–Wallis tests. NASA-TLX subscales and total workload scores were also analyzed using Kruskal–Wallis tests, as reported in Table 3. Outliers were identified as data points with standardized Z scores exceeding ± 3 and were handled using a Winsorizing approach [47]. All Kruskal–Wallis tests involved two independent groups ($df = 1$).

6.2 | Memory Outcomes, Diet Intake Attitudes, and Behavioral Commitment Intentions to Adopt Healthier Eating Habits

All three outcomes were analyzed using pre-post change scores, computed as the difference between post-interaction and pre-interaction measures. No significant modality differences were observed for memory outcomes and diet attitude between 2D and VR across all conditions. Behavioral commitment intentions to adopt healthier eating habits differed by modality only in the CON condition. Participants in the VR modality reported stronger behavioral commitment than those in the 2D modality, as indicated by a significant Kruskal–Wallis test, $H(1) = 4.25$, $p = .039$, with a small-to-moderate effect size ($\eta_H^2 = .086$).

6.3 | Working Alliance

We compared perceived working alliance between the 2D and VR modalities across the *Goal*, *Task*, *Bond*, and *Total Working Alliance* dimensions within each experimental condition.

Significant differences between 2D and VR modes emerged only within the CON condition (see Table 2). No significant differences were observed in the CSU, CVI, or CSV conditions.

6.4 | NASA TLX-Cognitive Load

NASA-TLX scores were compared between the 2D and VR modalities using rank-based Kruskal–Wallis tests conducted separately within each condition (CON, CSU, CVI, CSV). Table 3 reports means, standard deviations, Kruskal–Wallis test statistics (H), p -values, and eta-squared effect sizes (η_H^2). Because total NASA-TLX workload was computed as a normalized weighted composite, these values should be interpreted as relative differences in overall perceived workload rather than as raw 0 to 100 scores. The individual NASA-TLX subscales, however, were collected on the standard 0 to 100 scale and remain directly interpretable. Significant modality effects were observed for four of the six TLX subscales (*Mental Demand*, *Temporal Demand*, *Physical Demand*, and *Effort*), as well as for *Total Workload*.

6.5 | Perceptual Impressions of the Agent

Wilcoxon rank-sum tests were used to compare the agent impression measures between the 2D and VR modalities within each experimental condition. In the CON condition, 2D participants rated the agent more favorably, reporting greater comfort, appropriateness, and engagement, and less awkwardness and coldness. In the CSU condition, VR participants gave more negative impressions (eeriness, weirdness, unnaturalness, and dislike), while 2D participants found the agent more enjoyable. In the CVI condition, VR participants reported paying more attention, and in the CSV condition, they rated the agent as more energizing. Table 4 summarizes the significant comparisons with descriptive statistics and test results.

The results also indicated that CON and CSU elicited stronger subjective impressions compared to CVI and CSV. A possible explanation lies in the fact that in the CVI and CSV conditions, participants' gaze was directed away from the agent because of the visual elements in the interaction. This suggests reduced visual engagement or alignment with the ECA. This comparatively weaker performance of the CVI and CSV conditions may be explained through the split-attention effect in multimedia learning, where users divide visual attention across multiple spatially separated information sources, increasing extraneous cognitive load [48]. In the CVI and CSV conditions, participants were required to alternate gaze between the ECA and the visualization display positioned behind the agent, potentially reducing sustained attention toward the ECA itself. Prior eye-tracking studies have shown that additional on-screen visual elements can divert fixation patterns away from primary instructional content and increase attentional fragmentation [49].

7 | Discussion

7.1 | RQ1: Do Conversations in VR With the ECA Lead to Higher Perceived Workload Compared to Conversations in A 2D Platform?

The findings indicate that VR was associated with higher perceived workload than 2D across multiple NASA TLX dimensions, with effects varying by workload component and support condition. Participants in the VR modality reported higher

TABLE 2 | Working alliance differences between 2D and VR within each support condition. Values are item-level means with standard deviations in parentheses. Kruskal–Wallis tests compare 2D and VR within each condition for each WAI dimension.

Condition	Dimension	2D M (SD)	VR M (SD)	<i>H</i>	<i>p</i>
CON	Goal	3.92 (1.24)	3.04 (1.12)	6.55	0.011
	Task	3.90 (0.94)	2.99 (0.93)	6.94	0.008
	Bond	3.49 (1.30)	2.54 (1.03)	5.68	0.017
	Total WAI	3.77 (1.08)	2.85 (0.96)	6.83	0.009
CSU	Goal	4.00 (0.50)	3.56 (0.80)	3.66	0.056
	Task	3.61 (0.55)	3.46 (0.79)	0.30	0.585
	Bond	3.48 (0.90)	3.14 (0.99)	0.67	0.414
	Total WAI	3.70 (0.54)	3.39 (0.73)	1.66	0.198
CVI	Goal	3.84 (0.84)	3.89 (0.95)	0.23	0.634
	Task	3.29 (0.80)	3.46 (0.75)	0.71	0.399
	Bond	3.35 (0.80)	3.28 (1.02)	0.01	0.935
	Total WAI	3.49 (0.74)	3.54 (0.79)	0.09	0.766
CSV	Goal	3.96 (1.02)	3.78 (0.71)	1.12	0.289
	Task	3.70 (0.91)	3.61 (0.62)	1.04	0.307
	Bond	3.56 (1.09)	3.40 (0.80)	0.60	0.438
	Total WAI	3.74 (0.82)	3.60 (0.56)	0.75	0.386

TABLE 3 | NASA TLX subscale results comparing workload between 2D and VR within each condition (CON, CSU, CVI, CSV). Means (SD), Kruskal–Wallis test statistics (*H*), eta-squared effect sizes (η^2_H), and *p*-values are reported. Rows are shown only for comparisons with *p* < .05.

Condition	Subscale	2D Score	VR Score	<i>H</i>	η^2_H	<i>p</i>
CON	Physical	1.50 (2.95)	11.55 (15.51)	7.18	0.163	0.007
CSU	Physical	1.00 (1.72)	13.35 (14.90)	8.15	0.188	0.004
CVI	Mental	15.75 (14.63)	33.20 (25.33)	5.50	0.118	0.019
	Physical	2.20 (6.72)	13.30 (16.02)	4.19	0.084	0.041
	Total Workload	6.62 (2.76)	9.99 (3.71)	9.43	0.222	0.002
CSV	Mental	15.75 (16.46)	44.10 (21.01)	16.90	0.419	< .001
	Physical	1.25 (2.47)	6.50 (10.12)	3.94	0.077	0.047
	Temporal	12.20 (13.39)	25.95 (14.75)	8.51	0.198	0.004
	Effort	10.00 (10.59)	22.95 (18.74)	11.50	0.276	< 0.001
	Total workload	7.08 (2.17)	10.62 (2.73)	15.80	0.390	< 0.001

physical demand across conditions, while differences in mental demand, effort, temporal demand, and overall workload emerged primarily in conditions with additional multimodal support, particularly CSV (see Table 3). Importantly, these effects were not uniform across all subscales: no significant differences emerged for frustration or performance, suggesting that elevated workload reflected increased processing demands rather than perceptions of failure or inefficiency. Based on these findings, we partially support H1. Because total NASA-TLX workload was computed as a normalized weighted composite, observed differences should be interpreted as relative increases in perceived interaction effort rather than absolute workload levels.

Evidence from prior work suggests that differences between 2D and 3D simulation environments may partly reflect cognitive

overload and attentional distraction rather than representational fidelity alone. Richards and Taylor found that students achieved stronger learning outcomes with a 2D simulation than with a 3D virtual world when learning animal behavior concepts, arguing that the richer 3D representation may have introduced extraneous perceptual information that distracted learners from the core conceptual model [50]. This interpretation is relevant to the present findings. Although VR did not reduce memory retention in our study, it was associated with higher perceived workload, especially when visualizations and subtitles were added. The CVI and CSV conditions suggest that additional support features can become cognitively costly in immersive environments when users must divide attention between the agent, subtitles, visualizations, and the surrounding virtual space. These workload effects may also help explain why subjective impressions

TABLE 4 | Differences between 2D and VR by condition. The table reports mean (SD), Wilcoxon test statistics (W), effect sizes (r), and p -values for comparisons with $p < .05$.

Condition	Variable	2D Mean (SD)	VR Mean (SD)	W	Effect sizes (r)	p
CON	COMFORT	6.10 (1.55)	5.30 (1.56)	273.5	0.328	0.0378
	APPROP	5.65 (1.63)	4.10 (1.68)	302.5	0.445	0.0049
	ENGAGE	5.35 (1.84)	4.15 (1.66)	277.0	0.335	0.0339
	AWKWARD	2.20 (1.58)	3.20 (1.58)	121.5	0.343	0.0300
	COLD	2.40 (1.64)	3.90 (1.74)	109.5	0.394	0.0126
	LIKENESS	5.20 (1.79)	3.85 (1.63)	283.5	0.362	0.0222
	DISLIKENESS	2.35 (1.81)	3.75 (1.89)	108.0	0.400	0.0115
CSU	UNNAT	2.60 (1.82)	3.75 (1.68)	120.0	0.345	0.0289
	EERIE	2.25 (1.59)	3.50 (1.47)	104.0	0.424	0.0073
	WEIRD	2.00 (1.62)	2.85 (1.42)	113.0	0.383	0.0153
	FUN	5.20 (1.82)	3.85 (1.63)	289.0	0.385	0.0150
	AWKWARD	1.90 (1.48)	2.50 (1.10)	123.0	0.342	0.0304
	COLD	1.75 (1.41)	3.30 (1.63)	76.5	0.545	< 0.001
	DISLIKENESS	1.75 (1.41)	2.75 (1.45)	104.0	0.428	0.0068
CVI	IGNORED	1.65 (1.39)	2.20 (1.15)	121.5	0.362	0.0219
	ATTENTION	4.70 (1.34)	5.60 (1.50)	114.0	0.380	0.0162
CSV	ENERGIZED	3.10 (1.83)	4.30 (1.66)	123.5	0.329	0.0373

varied across support conditions rather than uniformly improving with added multimodal information. In contrast, the stronger working alliance observed in the 2D conversation-only condition suggests that lower perceptual complexity may support relational alignment when the interaction depends primarily on the agent's conversational behavior.

7.2 | RQ2: How do Conversations on A 2D Platform Versus in A VR Environment Affect the Formation of A Working Alliance with An Embodied Conversational Agent?

In the Conversation-Only (CON) condition, participants in the 2D modality reported significantly stronger working alliance across all dimensions, including goal alignment, task agreement, emotional bond, and overall alliance, compared to those in VR (See Table 2.) These differences did not appear in conditions with subtitles, visualizations, or their combination, where scores were comparable across modalities. This pattern suggests that modality effects on alliance were most pronounced in minimally supported interactions, whereas additional multimodal supports reduced relational differences across platforms. Based on these findings, we partially accept H2.

7.3 | Perceptual Impressions of the Agent

The perceptual impression results suggest that multimodal support did not uniformly improve how participants experienced the agent. Descriptively, the CON and CSU conditions showed more favorable subjective impression patterns than the CVI and CSV conditions. One possible explanation is that, in the CVI and

CSV conditions, participants' gaze was directed away from the agent because of the added visual elements in the interaction. As shown in Figure 3, these conditions required participants to divide visual attention between the ECA and the visualization display, which may have reduced sustained visual engagement with the agent itself. This interpretation is consistent with the split-attention principle in multimedia learning, where spatially separated information sources can increase extraneous cognitive load [48]. It is also consistent with eye-tracking evidence showing that additional visual elements can divert fixation patterns away from primary instructional content and increase attentional fragmentation [49]. In the present study, this attentional division may help explain why CVI and CSV produced less favorable subjective impression patterns despite providing additional informational support.

7.4 | RQ3: How do Conversations on A VR Platform Versus A 2D Platform Influence Participants' Behavioral Commitment Intentions to Adopt Healthier Eating Habits Following An ECA-Led Health Conversation?

Results showed a significant difference only in the Conversation-Only (CON) condition. Participants in the VR modality reported stronger commitments to adopting healthy eating behaviors compared to those in 2D. In conditions with subtitles, visualizations, or their combination, no significant differences between modalities were observed. Thus, the effect associated with VR emerged primarily in minimally supported interactions, while intentions were comparable across platforms when additional supports were present. Based on these findings, H3 was supported only in the CON condition.

7.5 | RQ4: Does the Use of A VR Platform Influence Memory Retention During Health-Related Conversations with An ECA Compared to A 2D Platform?

Across all conditions, memory retention scores did not significantly differ between 2D and VR modalities (see Section 6.2). Pre-post knowledge gains were comparable regardless of platform, indicating that immersive presentation did not systematically alter recall of health-related information in this context. H4 was not supported; no modality-dependent differences in memory retention were observed across multimodal support conditions.

8 | Conclusions

This study advances understanding of how immersion fidelity and multimodal support jointly shape extended interactions with embodied conversational agents [51]. Importantly, both platforms supported meaningful gains across core health-related outcomes, indicating that effectiveness can be achieved through different experiential pathways rather than a single optimal modality. Rather than treating immersive fidelity as a uniform pathway to improved outcomes, the findings demonstrate that 2D and VR platforms afford distinct interaction dynamics that differentially influence cognitive workload, relational alignment, and behavioral commitment intentions to adopt healthier eating habits in minimally supported, conversation-only interactions, depending on context and support structure.

A central contribution of this work is its focus on sustained, 20+ min health counseling conversations. By extending interaction length beyond the brief encounters typical of prior ECA studies, the study captures processes that unfold over time, including rapport formation, reflective engagement, and goal alignment. This temporal depth aligns digital health interactions more closely with established counseling and motivational interviewing practices, where change is understood as an emergent process rather than an immediate response to information delivery [15, 16].

Across conditions, immersive VR was associated with higher physical workload, with higher cognitive and overall workload emerging particularly when multiple multimodal supports were present. These effects are consistent with cognitive load accounts in which additional representational channels compete for limited processing resources, especially in immersive contexts that already impose increased perceptual and attentional demands [23, 24, 29]. At the same time, VR uniquely supported stronger behavioral commitment in minimally scaffolded, conversation-only interactions, specifically in the absence of additional multimodal supports. In contrast, 2D interactions supported lower workload and more stable relational alignment in these settings, facilitating emotional bond and task agreement without additional processing demands. Importantly, these patterns reflect differences in underlying mechanisms rather than evidence of platform superiority.

Although the present study did not measure downstream health behaviors directly, the observed differences in behavioral commitment intentions remain meaningful in light of prior work

demonstrating a positive association between attitudes, intentions, and subsequent behavior change across health domains [44, 45]. From this perspective, behavioral commitment intentions can be interpreted as proximal indicators of motivational readiness rather than immediate behavior change.

The asymmetric effects of subtitles and visualizations further extend multimedia learning theory into embodied conversational settings. While such supports often reinforce clarity and comprehension in non-immersive contexts [20, 21], they can increase workload or alter social impressions in immersive environments when they compete with embodied co-presence cues [23, 24]. These findings indicate that identical multimodal supports may operate differently depending on immersive context and interaction demands.

Taken together, the results motivate a design-oriented interpretation rather than a platform verdict. When relational stability, clarity, and sustained dialogue are central, lower-fidelity interfaces can provide a robust foundation for extended health conversations. When experiential salience and motivational activation are priorities, immersive VR can offer distinct benefits, even when accompanied by increased cognitive and physical demands. Hybrid approaches that leverage the complementary strengths of both platforms may therefore be especially promising for scalable health interventions [52].

In sum, immersion fidelity and multimodal support should be treated as interacting design dimensions that shape how cognitive effort, relational alignment, and behavioral commitment intentions emerge over time. Recognizing these tradeoffs is essential for the effective design of embodied conversational agents in health contexts.

9 | Limitations and Future Work

This study has several limitations. Although conversations with the agent averaged over 23 min, unusually long for virtual reality, multi-session studies are needed to examine sustained engagement, alliance formation, and changes in behavioral commitment intentions to adopt healthier eating habits over time. In our setup, the interaction was primarily frontal and all key elements, including the agent, subtitles, whiteboard, and TV, fit within the headset field of view, which likely reduced the need for large head rotations and extensive optical flow. In addition, the VR condition used an HTC Vive Pro Eye headset, and its display characteristics may differ from those of newer head-mounted displays with higher resolution. As a result, some perceptual, workload, or interaction outcomes may have been influenced in part by the specific hardware used in this study. Although the HTC Vive Pro Eye also supports integrated eye tracking, this study did not process gaze measures since it is out of the scope of this study. This represents an important direction for future work, particularly for examining split attention and visual cognitive load under subtitle and visualization conditions. Moreover, the 2D condition was presented on a 55-inch display to approximate a more life-sized interaction with the agent, based on prior peer-reviewed virtual human research using large-screen setups to support more realistic interpersonal encounters. At the same time, this design choice may have increased co-location or interpersonal salience in the

2D condition and may therefore have contributed to stronger working alliance responses. Future work should directly compare standard desktop displays, large-screen displays, and immersive VR during multimodal conversational interactions with embodied conversational agents to isolate how display size and immersive fidelity shape attention allocation, workload, and user experience. Future work should also directly measure head rotation and related movement patterns to better assess how these factors may shape user experience in VR [53, 54]. The study also focused on a single task domain and agent design, which may limit generalizability to other contexts or interaction styles.

Author Contributions

Mr. Sagar A. Vankit: contributed to the study design, data collection, data analysis, and manuscript preparation. **Mrs. Nafisa Sabnam Shama:** contributed to the study design and manuscript preparation. **Mr. Jay Ashish Kadam:** contributed to the study design and manuscript preparation. **Dr. Vivian Genaro Motti:** contributed to manuscript preparation. **Dr. Joseph LaViola:** contributed to manuscript preparation. **Dr. Andrew Duchowski:** contributed to manuscript preparation. **Dr. Matias Volonte:** led the study design, data analysis, and manuscript preparation.

Acknowledgments

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Funding

The authors have nothing to report.

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The authors have nothing to report.

References

- World Medical Association, "WMA Statement on the Ethics of Telemedicine," (2020), Accessed: 2025-08-25, <https://www.wma.net/policies-post/wma-statement-on-the-ethics-of-telemedicine/>.
- Group, Diabetes Prevention Program (DPP) Research, "The Diabetes Prevention Program (DPP) Description of Lifestyle Intervention," *Diabetes Care* 25, no. 12 (2002): 2165–2171.
- D. T. Michaeli, J. C. Michaeli, S. Albers, and T. Michaeli, "The Healthcare Workforce Shortage of Nurses and Physicians: Practice, Theory, Evidence, and Ways Forward," *Policy, Politics & Nursing Practice* 25, no. 4 (2024): 216–227.
- G. Clarke and B. J. Yarborough, "Evaluating the Promise of Health IT to Enhance/Expand the Reach of Mental Health Services," *General Hospital Psychiatry* 35, no. 4 (2013): 339–344.
- T. Bickmore, A. Gruber, and R. Picard, "Establishing the Computer–Patient Working Alliance in Automated Health Behavior Change Interventions," *Patient Education and Counseling* 59, no. 1 (2005): 21–30.
- L. Laranjo, A. G. Dunn, and H. L. Tong, "Conversational Agents in Healthcare: A Systematic Review," *Journal of the American Medical Informatics Association* 25, no. 9 (2018): 1248–1258.
- J. Mercado, I. E. Espinosa-Curiel, and J. Martínez-Miranda, "Embodied Conversational Agents Providing Motivational Interviewing to Improve Health-Related Behaviors: Scoping Review," *Journal of Medical Internet Research* 25 (2023): e52097.
- D. Schulman, "Embodied Agents for Long-Term Discourse," PhD Thesis, Northeastern University, Boston, MA, (2013).
- T. Kowatsch, T. Schachner, S. Harperink, et al., "Conversational Agents as Mediating Social Actors in Chronic Disease Management Involving Health Care Professionals, Patients, and Family Members: Multisite Single-Arm Feasibility Study," *Journal of Medical Internet Research* 23, no. 2 (2021): e25060.
- Z. Jiang, X. Huang, Z. Wang, Y. Liu, L. Huang, and X. Luo, "Embodied Conversational Agents for Chronic Diseases: Scoping Review," *Journal of Medical Internet Research* 26 (2024): e47134.
- L. L. Kramer, S. Ter Stal, B. C. Mulder, E. de Vet, and L. van Velsen, "Developing Embodied Conversational Agents for Coaching People in a Healthy Lifestyle: Scoping Review," *Journal of Medical Internet Research* 22, no. 2 (2020): e14058.
- T. Bickmore, D. Schulman, and L. Yin, "Maintaining Engagement in Long-Term Interventions With Relational Agents," *Applied Artificial Intelligence* 24, no. 6 (2010): 648–666.
- A. H. Berman, M. Beckman, and H. Lindqvist, "Motivational Interviewing Interventions," in *Cambridge Handbooks in Psychology* (Cambridge University Press, 2020), 661–676.
- C. C. Butler, S. Rollnick, D. Cohen, M. Bachmann, I. Russell, and N. Stott, "Motivational Consulting Versus Brief Advice for Smokers in General Practice: A Randomized Trial," *British Journal of General Practice* 49, no. 445 (1999): 611.
- W. R. Miller and S. Rollnick, *Motivational Interviewing: Helping People Change* (Guilford press, 2012).
- S. Rollnick, W. R. Miller, C. C. Butler, and M. S. Aloia, "Motivational Interviewing in Health Care: Helping Patients Change Behavior," (2008).
- S. Baptista, G. Wadley, D. Bird, et al., "Acceptability of an Embodied Conversational Agent for Type 2 Diabetes Self-Management Education and Support via a Smartphone App: Mixed Methods Study," *JMIR mHealth and uHealth* 8, no. 7 (2020): e17038.
- S. Lim, R. Schmälzle, and G. Bente, "Artificial Social Influence via Human-Embodied AI Agent Interaction in Immersive Virtual Reality (VR): Effects of Similarity-Matching During Health Conversations," *Computers in Human Behavior: Artificial Humans* 5 (2025): 100172.
- R. E. Mayer, *The Cambridge Handbook of Multimedia Learning* (Cambridge university press, 2005).
- J.-L. Kruger, "Subtitles in the Classroom: Balancing the Benefits of Dual Coding With the Cost of Increased Cognitive Load," *Journal for Language Teaching= Ijenali Yekufundzisa Lulwimi= Tydskrif Vir Taalonderrig* 47, no. 1 (2013): 29–53.
- P. S. Houts, C. C. Doak, L. G. Doak, and M. J. Loscalzo, "The Role of Pictures in Improving Health Communication: A Review of Research on Attention, Comprehension, Recall, and Adherence," *Patient Education and Counseling* 61, no. 2 (2006): 173–190.
- B. Liu, G. Salvendy, and T. Kuczek, "The Role of Visualization in Understanding Abstract Concepts," *International Journal of Cognitive Ergonomics* 3, no. 4 (1999): 289–305.
- R. E. Mayer and R. Moreno, "A Split-Attention Effect in Multimedia Learning: Evidence for Dual Processing Systems in Working Memory," *Journal of Educational Psychology* 90, no. 2 (1998): 312.
- J. Sweller, "Cognitive Load Theory, Learning Difficulty, and Instructional Design," *Learning and Instruction* 4, no. 4 (1994): 295–312.

25. G. Makransky, T. S. Terkildsen, and R. E. Mayer, "Adding Immersive Virtual Reality to a Science Lab Simulation Causes More Presence but Less Learning," *Learning and Instruction* 60 (2019): 225–236.
26. J. Parong and R. E. Mayer, "Learning Science in Immersive Virtual Reality," *Journal of Educational Psychology* 110, no. 6 (2018): 785.
27. U. Malinowska, J. Wojciechowski, M. Waligóra, and J. Rogala, "Corrigendum: Performance of Game Sessions in VR vs Standard 2D Monitor Environment. An EEG Study," *Frontiers in Physiology* 16 (2025): 1616912.
28. N. Wenk, J. Penalver-Andres, K. A. Buetler, T. Nef, R. M. Müri, and L. Marchal-Crespo, "Effect of Immersive Visualization Technologies on Cognitive Load, Motivation, Usability, and Embodiment," *Virtual Reality* 27, no. 1 (2023): 307–331.
29. J. J. Cummings and J. N. Bailenson, "How Immersive Is Enough? A Meta-Analysis of the Effect of Immersive Technology on User Presence," *Media Psychology* 19, no. 2 (2016): 272–309.
30. S. H. Yip and J. A. Saunders, "Restricting the Distribution of Visual Attention Reduces Cybersickness," *Cognitive Research: Principles and Implications* 8, no. 1 (2023): 18.
31. R. C. A. Barrett, R. Poe, J. W. O'Camb, et al., "Comparing Virtual Reality, Desktop-Based 3D, and 2D Versions of a Category Learning Experiment," *PLoS One* 17, no. 10 (2022): e0275119.
32. S. A. Vankit, V. G. Motti, T. D. Do, et al., "Path Modeling of Visual Attention, User Perceptions, and Behavior Change Intentions in Conversations With Embodied Agents in VR," *Computer Animation and Virtual Worlds* 36, no. 3 (2025): e70028.
33. E. Yorgancigil, B. A. Urgan, and F. Yildirim, "Uncanny Valley Effect is Amplified with Multimodal Stimuli and Varies Across Ages," (2021).
34. J. Gratch, "The Promise and Peril of Interactive Embodied Agents for Studying Non-Verbal Communication: A Machine Learning Perspective," *Philosophical Transactions of the Royal Society, B: Biological Sciences* 1875 (2023): 2021047.
35. C. Flückiger, A. C. Del Re, B. E. Wampold, and A. O. Horvath, "The Alliance in Adult Psychotherapy: A Meta-Analytic Synthesis," *Psychotherapy* 55, no. 4 (2018): 316.
36. M. Volonte, A. Robb, A. T. Duchowski, and S. V. Babu, "Empirical Evaluation of Virtual Human Conversational and Affective Animations on Visual Attention in Inter-Personal Simulations," in *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, 25–32. (IEEE, 2018).
37. T. Bickmore, H. Trinh, R. Asadi, and S. Olafsson, "Safety First: Conversational Agents for Health Care," *Studies in Conversational UX Design* (2018): 33–57.
38. P. Ekman and W. V. Friesen, "Facial Action Coding System," *Environmental Psychology & Nonverbal Behavior* 2 (1978).
39. M. Volonte, E. Ofek, K. Jakubzak, S. Bruner, and M. Gonzalez-Franco, "Headbox: A Facial Blendshape Animation Toolkit for the Microsoft Rocketbox Library," in *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, 39–42. (IEEE, 2022).
40. N. Aburumman, M. Gillies, J. A. Ward, and A. F. de C Hamilton, "Non-verbal Communication in Virtual Reality: Nodding as a Social Signal in Virtual Interactions," *International Journal of Human-Computer Studies* 164 (2022): 102819.
41. C. Kühne, E. D. Kecelioglu, S. Maltby, and M. G. Kluge, "Direct Comparison of Virtual Reality and 2D Delivery on Sense of Presence and Emotional Arousal," *Frontiers in Virtual Reality* 4 (2023): 1211001.
42. U.S. Department of Agriculture, "What Is MyPlate?," accessed February 5, 2026, [MyPlate.gov](https://www.mypplate.gov).
43. B. Lundahl, T. Moleni, B. L. Burke, et al., "A Meta-Analysis of Motivational Interviewing: Twenty-Five Years of Empirical Studies," *Research on Social Work Practice* 23, no. 2 (2013): 137–160.
44. I. Ajzen, "The Theory of Planned Behavior," *Organizational Behavior and Human Decision Processes* 50, no. 2 (1991): 179–211.
45. P. Sheeran, "Intention–Behavior Relations: A Conceptual and Empirical Review," *European Review of Social Psychology* 12, no. 1 (2002): 1–36.
46. A. O. Horvath and L. S. Greenberg, *The Working Alliance: Theory, Research, and Practice* (John Wiley & Sons, 1994).
47. D. M. Erceg-Hurn and V. M. Mirosevich, "Modern Robust Statistical Methods: An Easy Way to Maximize the Accuracy and Power of Your Research," *American Psychologist* 63, no. 7 (2008): 591.
48. P. Ayres and J. Sweller, "The Split-Attention Principle in Multimedia Learning," in *The Cambridge Handbook of Multimedia Learning* 2 (2005), 135–146.
49. D. Mutlu-Bayraktar, P. Ozel, F. Altindis, and B. Yilmaz, "Split-Attention Effects in Multimedia Learning Environments: Eye-Tracking and EEG Analysis," *Multimedia Tools and Applications* 81, no. 6 (2022): 8259–8282.
50. D. Richards and M. Taylor, "A Comparison of Learning Gains When Using a 2D Simulation Tool Versus a 3D Virtual World: An Experiment to Find the Right Representation Involving the Marginal Value Theorem," *Computers & Education* 86 (2015): 157–171.
51. S. A. Vankit, S. Zamanifard, D. Diaz, et al., "Exploring the Impact of Multimodal Long Conversations in VR on Attitudes Toward Behavior Change, Memory Retention, and Cognitive Load," *Computer Animation and Virtual Worlds* 36, no. 3 (2025): e70023.
52. F. De Lorenzis, A. Visconti, S. Restivo, et al., "Combining Virtual Reality With Asymmetric Collaborative Learning: A Case Study in Chemistry Education," *Smart Learning Environments* 11, no. 1 (2024): 43.
53. S. Palmisano, R. Mursic, and J. Kim, "Vection and Cybersickness Generated by Head-And-Display Motion in the Oculus Rift," *Displays* 46 (2017): 1–8.
54. N. Tian, P. Lopes, and R. Boulic, "A Review of Cybersickness in Head-Mounted Displays: Raising Attention to Individual Susceptibility," *Virtual Reality* 26, no. 4 (2022): 1409–1441.