

The Impact of Health Literacy in a Virtual Agent Intervention Addressing Barriers to Cancer Clinical Trial Participation

Anonymous Author(s)

Abstract

Communication barriers contribute low clinical trial participation, which is crucial in advancing breast cancer treatment. Patient-centered communication and considering user characteristics like health literacy may improve communication about clinical trials. Virtual agent interventions can employ these strategies to improve communication. However, it is unclear how the adoption of these strategies impact trust and use of a virtual agent intervention. Therefore, we conducted a study where 103 breast cancer patients interacted with a web-based patient-centered virtual agent intervention to learn about clinical trials. Although the virtual agent's communication strategy alone did not influence perceptions of trust, findings revealed that accounting for participants' health literacy significantly impacted trust and unveiled a more nuanced understanding of patient-centered communication strategies. This work contributes to the design of virtual agent interventions, highlighting the necessity of incorporating user characteristics like health literacy when developing interventions for breast cancer patients for clinical trial engagement.

CCS Concepts

• Human-centered computing → User studies.

Keywords

virtual agents, user characteristics, healthcare interventions, clinical trials

ACM Reference Format:

Anonymous Author(s). 2018. The Impact of Health Literacy in a Virtual Agent Intervention Addressing Barriers to Cancer Clinical Trial Participation. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation email (Conference acronym 'XX)*. ACM, New York, NY, USA, 11 pages. <https://doi.org/XXXXXXX.XXXXXXX>

1 Introduction

Breast cancer is one of the most common cancers [62], making up 16% of new cancer cases in 2024 [50]. Advances in cancer treatments are largely driven by clinical trials and participation from cancer patients. However, adult participation in cancer clinical trials remains low, ranging from 2% to 8% [76, 77]. Barriers to enrollment include a general lack of knowledge and awareness of clinical trials

[39, 45, 75], and mistrust and stigma towards medical establishments and research [4, 32].

Online educational interventions, such as videos [63] and e-delivered learning modules [35, 61], have become a popular way to deliver health information. A growing approach in this space involves virtual agents – computer-generated characters that appear and behave like real humans [84]. Virtual agents can engage users through interactive, personalized conversations, which may enhance user understanding [11], trust [86], and intent to act on health information [88]. The Computers Are Social Actors Framework (CASA) [49, 58] suggests that virtual agents are perceived through a social lens, allowing for strategic tailoring of features such as age, gender, personality, and communication strategy [1, 46, 74]. To this end, virtual agents' accessibility and adaptability make them effective for promoting healthy behaviors across diverse users and contexts [66, 89, 91].

Literature suggests that improving communication about clinical trials may address barriers to clinical trial engagement [4, 45], such as building legitimacy and trust [85]. In healthcare, there is an emphasis on patient-centered communication to foster trust [33, 95] and improve patient outcomes [30, 44, 47, 79], involving strategies such as using familiar language and expressing empathy when explaining health topics [23, 83, 94]. Regarding virtual agents, there is evidence that employing patient-centered communication strategies can improve health outcomes [90]. Therefore, applying these strategies in virtual agents holds promise for fostering trust and enhancing engagement with clinical trials.

Additionally, user characteristics (e.g., age and personality) may influence users' engagement with virtual agents [1, 46, 69, 74], though their impact varies by context [59]. Healthcare literature highlights health literacy as an important characteristic to consider in cancer patients [34] and engaging underrepresented populations in clinical trials who are disproportionately affected by low health literacy [22]. Prior work suggests that virtual agents can positively influence outcomes with low health literacy cancer patients for clinical trial-related interventions [10]. Hence, health literacy is a key user characteristic to consider for our context of a virtual agent promoting clinical trial engagement for breast cancer patients.

Therefore, this explores the use of patient-centered communication strategies in a virtual agent and the role of users' health literacy in fostering trust and promoting clinical trial engagement among breast cancer patients. We address the following research questions:

- **RQ1:** How does a virtual agent using patient-centered communication strategies influence perceived trust?
 - **RQ1a:** How does a user's engagement and health literacy moderate the perceived trust of virtual agent using patient-centered communication strategies?

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

Conference acronym 'XX, Woodstock, NY

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-XXXX-X/2018/06

<https://doi.org/XXXXXXX.XXXXXXX>

- **RQ2:** How does a virtual agent using patient-centered communication strategies influence a user's browse behavior for clinical trials?
 - **RQ2a:** How does a user's engagement and health literacy moderate their browse behavior when interacting with a virtual agent using patient-centered communication strategies?

2 Related Works

In this section, we provide an overview of the importance of trust, communication strategies, and user characteristics in health-related virtual agent interventions.

2.1 Trust in Health-Related Virtual Agent Interventions

Virtual agents are increasingly used in healthcare interventions such as screening, promoting healthy behaviors, and mental health interventions [12, 25, 89]. Virtual agents have also shown promise in clinical trial contexts by providing interactive, engaging interventions to address mistrust and misunderstanding [48]. As virtual agents take on more prominent roles in sensitive health domains, perceived trust becomes critical to user acceptance and adherence to their advice [13, 38, 51, 78]. This is particularly important in clinical trial engagement efforts, where behavioral recruitment efforts emphasize trust building [85]. Prior work suggests various factors impact trust [86], such as the agent's appearance, communication style, and the user's own characteristics, though which factors shape trust varies by context [20, 28, 80].

Therefore, despite the potential for virtual agent interventions for clinical trial engagement, it remains unclear what factors shape trust. Understanding what contributes to perceptions of trust can improve acceptance of virtual agent interventions and user outcomes. Therefore, this work contributes to the design of health-related virtual agent interventions by exploring how the virtual agent's communication style and users' characteristics impact perceptions of trust for a clinical trial awareness intervention.

2.2 Communication Strategies in Health-Related Virtual Agent Interventions

Prior work shows that the communication strategy employed by a virtual agent impacts agent perceptions and user outcomes in health-related interventions [1, 11, 88]. For example, empathetic dialogue can improve perceptions of virtual agents [2, 73], and persuasive dialogue and humor can improve user outcomes such as attitudes and motivation towards healthy habits [53, 60]. In healthcare literature, there is an emphasis on patient-centered communication which embodies strategies such as responding to patients' emotions and understanding the patients' perspectives, needs, and values [37, 41]. Patient-centered communication is positively associated with trust and may lead to better health outcomes [55] such as satisfaction with care, understanding medical information and recommendations, and adherence to therapy [57]. In practice, patient-centered communication includes linguistic strategies such as aligning communication using similar language, expressing empathy, and adjusting health literacy [23, 83, 94].

However, employing patient-centered communication strategies in virtual agent health-related interventions remains underexplored. Despite its wide use in healthcare, it is unclear how patient-centered communication may influence trust in virtual agents and user outcomes. Therefore, our work contributes to the design of virtual agent communication by exploring the implementation of patient-centered communication strategies.

2.3 User Characteristics and Health-Related Virtual Agent Interventions

Prior work suggests that user characteristics – such as age, personality, and self-efficacy [1, 46, 69, 87, 88] – play a role in how people engage with virtual agents. In cancer-related virtual agent interventions, characteristics such as rural identity, racial-ethnic background, experience with healthcare, and age have been shown to impact perceptions of virtual agents [18, 81]. One important user characteristic in healthcare contexts is health literacy [34]. Health literacy is particularly important for individuals from underrepresented groups [22], whose participation in cancer clinical trials is crucial to advancing cancer treatment for groups disproportionately affected by cancer. Virtual agents have the potential to address low health literacy by being easier to use, more satisfactory, and more useful for low health literacy populations [7, 10, 82]. Moreover, prior work has shown the outcomes for an adaptive virtual agent is affected by the health literacy of the user [21].

Therefore, this work aims to expand the current understanding of user health literacy and its influence on virtual agent interventions. Specifically, we explore how a participant's health literacy impacts their engagement and perceptions of a virtual agent intervention to raise awareness about clinical trials.

3 Methodology

This study examines how patient-centered communication influences trust in a virtual agent and intervention outcomes, and how user health literacy and engagement impacts these relationships and outcomes. To study these effects, we conducted a between-participants study where participants interacted with a virtual agent for clinical trial education through open-ended input. We manipulated the communication style of the virtual agent: **patient-centered** or **control**. We describe the implementation in detail in this section.

3.1 System and Agent Design

3.1.1 Virtual Agent Interface. The virtual agent system was a NodeJS application deployed via Amazon Web Services. The virtual agent was created using Synthesia [72], an online platform that generates realistic virtual characters. OpenAI's Audio API was used for real-time text-to-speech, using the voice *Nova*. The same virtual character was pilot-tested and administered across both conditions and was displayed as a chatbot-style interface (Figure 1). A list of topics to cover based on the top barriers to clinical trial participation for breast cancer patients [4, 32, 39, 45, 75] was displayed on the left of the chat area. For each response, the generated text appeared on the interface alongside audio playback, while the agent was visually animated with a glow effect and subtle head and mouth movements to cue activity.

3.1.2 *Agent Dialogue Generation.* To generate responses, the virtual human utilized OpenAI’s Assistants API [54] using a vector store. Several sources of information from reputable institutions (i.e., National Institute of Health, National Cancer Institute, Food and Drug Administration, and US Department of Health and Human Services) were stored in the vector store to provide a database that the virtual agent could use in its response. The vector store served as the knowledge base for the virtual agent’s answering capabilities. For any other user queries expanded beyond the scope of the knowledge base, the agent informed the user that they were incapable of providing an answer at this time.

Both the *control* and *patient-centered* virtual agents generated responses through persona-informed LLMs [71] [5]. In the control condition, the agent was prompted to simply respond to users’ queries with appropriate answers. In the patient-centered condition, the agent was provided additional language manipulations to adjust its responses based on three patient-centered communication strategies [37, 41]: (1) *empathy* - expressing empathy and reassurance, (2) *linguistic similarity* - using similar language as the user, and (3) *readability* - lowering the health literacy of the information.

The prompts were constructed using a framework adapted from Steenstra et al.[67], which integrates established prompt engineering practices[3] with a health counseling dialogue framework[9]. Specifically, we defined the system’s identity, goals, and constraints in both control and patient-centered conditions by incorporating elements such as persona, context, disambiguation, and wording[3]. For the patient-centered condition, we further applied best practices from patient-centered communication using the theory model[9]. Table 1 details the prompt segments, which condition(s) they appeared in, and corresponding themes from the prompt framework [67]. Outputs were validated by health communication experts. The following is an example of a control and patient-centered response to the question, "Are clinical trials safe?":

- **Control:** "Clinical trials are designed with safety as a priority. They undergo rigorous review and monitoring to ensure participant safety. However, like any medical intervention, they may carry risks, which are carefully explained to participants before enrollment."
- **Patient-Centered:** "Yes, clinical trials are designed with safety in mind. They follow strict guidelines to protect participants. Before joining, you’ll be informed about potential risks and benefits. It’s normal to feel a bit anxious, but know that safety is a top priority in these studies. If you have more questions or concerns, I’m here to help!"

3.1.3 *Linguistic Validation of Agent Dialogue.* In addition to expert validation of agent responses, we analyzed post-study transcripts to confirm the patient-centered condition by examining linguistic features of our three manipulations: empathy, linguistic similarity, and readability. For each participant, all virtual agent messages were combined into a single text, as were all user messages. Empathy and linguistic similarity were analyzed using Linguistic Inquiry and Word Count (LIWC) [43], a validated tool for evaluating language in patient-centered communication [23, 94]. Empathy was measured by the frequency of affect, emotion, and prosocial behavior words [6]. Linguistic similarity was evaluated with Language

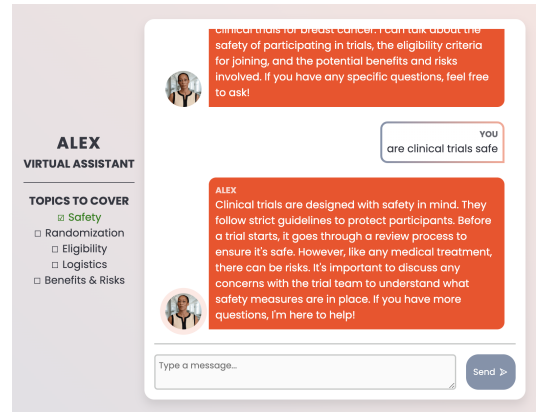


Figure 1: The patient-centered virtual agent intervention web interface.

Style Matching (LSM) [42], which computes stylistic alignment between user and agent text. Reading ease was calculated using the Flesch-Kincaid Grade Level (FKGL) index [24, 64], which estimates the education level required to understand given text. The calculated scores were compared between conditions to determine that the manipulated properties were improved in the patient-centered communication.

Empathy. The patient-centered virtual agent used significantly more affect ($U = 397.00, p < .001$), emotion ($U = 819.00, p < .001$), and prosocial behavior words ($U = 565.00, p < .001$) compared to the control, validating the empathy manipulation.

Linguistic Similarity. The patient-centered virtual agent’s responses demonstrated significantly higher language similarity between the virtual agent’s and user’s responses compared to the control ($U = 916.00, p = 0.003$), validating the linguistic similarity manipulation.

Readability. The patient-centered condition had a significantly lower FKGL at around a 10th-grade reading level ($U = 2764.00, p < 0.001$), compared to the control condition at >12th-grade, validating the readability manipulation. Although a 10th-grade reading level is considered high, the readability of the patient-centered condition is still lower than that of most clinical trial materials [24, 64].

3.2 Intervention Phases

The intervention consisted of three primary phases: Introduction, Clinical Trial Conversation, and Browse Studies. All users engaged in both the Introduction and Clinical Trial Conversation phases regardless of the condition to which they were assigned. Users, however, had the option to decide as to whether or not they would engage in the Browse Studies phase. Each phase is described in further detail in the following section.

Introduction. Once the user completed the IRB and pre-survey, they were randomized into either the *patient-centered* or *control* condition and were provided instructions on how to interact with the virtual agent. Participants were instructed to engage with the virtual agent through textual input and participate in at least five conversational turns pertaining to a set of suggested clinical trial

| Prompt Segment | Condition | Framework |
|---|------------------------------|---|
| <i>You are a virtual assistant educating on clinical trials for people with breast cancer.</i> | Control; Patient-Centered | Persona, Context [3] |
| <i>Use the Knowledge Base exclusively to answer questions.</i> | Control; Patient-Centered | Disambiguation, Context [3]; External Model [9] |
| <i>If an answer isn't there, inform the user you can't answer. Keep your response to 75 words or less.</i> | Control; Patient-Centered | Wording [3]; Protocol Model [9] |
| <i>Match your language style to the user: Mimic the user's tone, sentence structure, word choice, and casualness/formality. Adjust your health literacy based on the BRIEF SCORE (0-12, low health literacy; 13-20, medium to high health literacy). Express empathy and reassurance: Ease worries and anxieties. Acknowledge/validate the user's emotions/feelings. Use encouraging/supportive language.</i> | Patient-Centered | Theory Model [9] |

Table 1: Overview of prompt segments used in the study, illustrating which experimental condition(s) each snippet appeared in and the corresponding themes from the prompt framework adapted from [67]

topics. These topics were predetermined and created based on identified barriers to clinical trial participation [4, 32, 39, 45, 75]. Once the user finished reading the instructions, they began the Clinical Trial Conversation.

Clinical Trial Conversation. The primary phase of this intervention was the Clinical Trial Conversation. Participants engaged for a minimum of five conversational turns but were allowed to engage for as many conversational turns as desired. In the *control* condition, the virtual agent responded to the participant's queries based on its knowledge base. The *patient-centered* virtual agent responded with additional language manipulations based on patient-centered communication strategies (see Section 3.1). The *patient-centered* virtual agent checked users' understanding, suggested relevant next topics, and use empathetic language in addition to answering the participant's queries. After completing a minimum of five conversational turns, the participant was allowed to progress.

Browse Studies. The Browse Studies phase was the final phase in the intervention before the post-survey, and was optional. After completing the Clinical Trial Conversation with the virtual agent, participants were prompted with a choice to browse existing clinical trials on breast cancer. Participants who opted not to browse were immediately re-directed to the post-survey. Participants who opted to browse were shown three pre-selected breast cancer clinical trials with either the *control* or *patient-centered* virtual agent, based on the condition they were in. All the clinical trials that were presented to participants allowed for remote participation and were sourced from ClinicalTrials.gov, an online public database that contains information about clinical research studies. The virtual agent responded to any questions that the participant may have had about the pre-selected studies, where the *patient-centered* agent employed additional patient-centered communication strategies. Participants that chose to browse through the clinical research studies were allowed to move onto the post-survey at any point.

3.3 Measures

3.3.1 Outcome Measures.

Human-Computer Trust Scale. Participants answered twelve items about their trust [27] in the virtual character in the post-survey. These items captured four dimensions of trust, with three items each: risk perception (eg, "It is risky to interact with Alex"), benevolence (eg, "I believe that Alex will act in my best interest), competence (eg, "If I use Alex, I think I would be able to depend on it completely"), and reciprocity (eg, "I can trust the information presented to me by Alex"). Each item was scored on a 5-point Likert scale where 1 was "Strongly Disagree" and 5 was "Strongly Agree," with the risk perception items being reverse scored.

Average Words Per Dialogue Turn. Literature shows that patient participation reflects an overall orientation to active and engaged communication [68, 70]. We chose to measure participation by average words per dialogue turn in order to determine engagement with the intervention. Average words per dialogue turn was chosen over number of dialogue turns, as participants were required a set amount of dialogue turns for the intervention. Hence, average words per dialogue turn was a more informative indication of participant engagement.

Browse Choice. After interacting with the virtual agent, participants were asked if they would like to browse pre-selected clinical trials. Their response was recorded as "yes" or "no."

Follow Up Choice. For participants that chose to browse through the pre-selected clinical trials, they were asked if they would like follow up information. Their response was recorded as "yes" or "no."

Health Literacy. Participants' health literacy was measured in the pre-survey using the BRIEF Health Literacy Screening Tool [29]. This tool consists of 4 questions (eg, "How often do you have someone help you read hospital materials?"). Three of these questions are scored on a 5 point Likert scale where 1 = "Always" and 5 = "Never". One question ("How confident are you filling out medical forms by yourself?") is reverse scored, where 1 = "Not at all" and 5 = "Extremely." All four items were summed for a total score. Participants with a final score between 4-12 were grouped as *low*

health literacy (n=27), 13-16 were grouped as *medium health literacy* (n=31), and 17-20 were grouped as *high health literacy* (n=44).

3.4 Procedure

Participants were recruited online through Prolific and were provided with a link to complete the pre-survey. Then, participants were randomized to an intervention condition. Participants were not informed of their assigned condition but were told that the agent was a virtual assistant for educating on clinical trials in relation to breast cancer. After finishing the intervention, participants were redirected to completing the post-survey. Figure 2 provides an overview of the procedure described in this section.

Pre-Survey. Participants were redirected to the pre-survey on Qualtrics, which began with the informed consent. Then, participants responded to 4 questions from the BRIEF Health Literacy Screening Tool [29] in order to determine their *health literacy*. This BRIEF score was also used to inform the *patient-centered* virtual agent.

Intervention. After completing the pre-survey, participants were randomized to the *control* or *patient-centered* condition. They were redirected to the intervention website and completed each of the Intervention Tasks described in Section 3.2. After completing the entirety of the virtual agent interaction, participants were redirected to the post-survey.

Post-Survey. The post-survey contained the survey items for the outcome measures used in this study as well as demographics. Furthermore, if a participant had agreed to browse the sample studies, a question was provided in the survey asking if the participant wanted to follow up with additional information on clinical trials. Users who agreed to follow up were provided with clinicaltrials.gov links to the pre-selected clinical trials [15–17]. Participants who did not browse for clinical trials were not presented with this item.

3.5 Participants

This study was approved by the university’s Institutional Review Board (IRB#ET00042394). Participants (N = 103) were recruited through Prolific and were eligible if residing in the United States and reporting a current or past diagnosis of breast cancer. Each participant received monetary compensation of 5.00 USD upon study completion. The population’s reported age ranged from 21 to 79 years old (M = 50.62, SD = 15.19). Participants’ gender was reported as follows: 93.2% female (n=96), 5.8% male (n=6), 1.0% agender (n=1). The skewed groups in gender are expected given the recruitment requirements for a breast cancer diagnosis. Racial ethnic groups were reported as follows: 6.8% of our participants were Asian (n=7), 16.5% Black/African American (n=17), 1.0% Hispanic/Latin American (n=1), 71.8% White: Non-Hispanic, Non-Latin American (n=74), and 1.9% Middle Eastern/North African (n=2). Two participants did not report their racial ethnic group.

4 Results

Data analysis was conducted in IBM SPSS Statistics 30. We aimed to explore the influence of the *virtual agent’s communication style* on *trust* and *browse behavior*, and the effects that *words per dialogue turn* and *user health literacy* would have on our outcomes.

Therefore, we evaluated trust outcomes with hierarchical multiple regression, and browse behavior outcomes with hierarchical binomial regression. Internal reliability for the human-computer trust scale was assessed by calculating Cronbach’s alpha for the 4 subscales: risk perception ($\alpha = 0.794$), benevolence ($\alpha = 0.776$), competence ($\alpha = 0.840$), and reciprocity ($\alpha = 0.790$), as well as for all 12 trust items combined ($\alpha = 0.804$), indicating good to excellent internal reliability.

Assumptions for regression were checked prior to each analysis. Hierarchical models were built by entering Condition first, followed by Words per Dialogue Turns and interaction effects, and finally Low Health Literacy and interaction effects. We did not include Medium and High Health Literacy levels as they did not significantly improve model fit for any of our analyses. If a main effect of Low Health Literacy was found, post-hoc analysis was conducted to compare it to Medium and High Health Literacy.

4.1 Trust.

Risk Perception. The final model for Risk Perception was significant (see Table 2 for model steps). Significant main effects emerged for Condition, Words per Dialogue Turns, and Health Literacy. Participants in the patient-centered condition reported higher risk perception scores compared to those in the control condition. Participants who had higher words per dialogue turn were also associated with higher risk perception scores. For health literacy, post hoc comparisons using estimated marginal means confirmed that low health literacy participants ($M = 10.30, SD = 3.75; EMM = 10.34, SE = 0.56$) reported significantly lower risk perception scores than medium and high literacy participants ($M = 7.05, SD = 2.52; EMM = 7.03, SE = 0.33$), after adjusting for Words per Dialogue Turn. Finally, a significant interaction between Condition and Words per Dialogue Turn was also observed, $t(95) = 2.497, p = 0.014, r = 0.248$ (see Figure ??, revealing that as Words per Dialogue Turn increased, participants in the patient-centered condition had higher Risk Perception scores compared to the control condition.

| | Model 1 | Model 2 | Model 3 |
|------------------------------|----------------|----------------|----------------|
| (Intercept) | 8.00 (0.45)*** | 9.96 (0.88)*** | 8.97 (0.85)*** |
| Condition (Patient-Centered) | -0.16 (0.64) | -1.99 (1.29) | -3.05 (1.31)* |
| Words per Dialogue Turn | | -0.20 (0.08)* | -0.17 (0.07)* |
| Condition × Words | | 0.18 (0.11) | 0.26 (0.11)* |
| Health Literacy (Low) | | | 2.77 (0.89)** |
| Condition × Health Literacy | | | 1.06 (1.32) |
| R ² | 0.001 | 0.065 | 0.261 |
| Adj. R ² | -0.009 | 0.036 | 0.222 |
| F Change | 0.064 | 3.341* | 12.578*** |
| F | 0.064 | 2.250 | 6.703*** |

Table 2: Multiple regression model for Risk Perception with β (standard error). F-statistic with regard to individual model; F Change with regard to prior model step. * = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.05$**

Benevolence. The final model for Benevolence was significant (see Table 3 for model steps). Significant main effects emerged only for Health Literacy. Post hoc comparisons using estimated marginal means confirmed that low health literacy participants

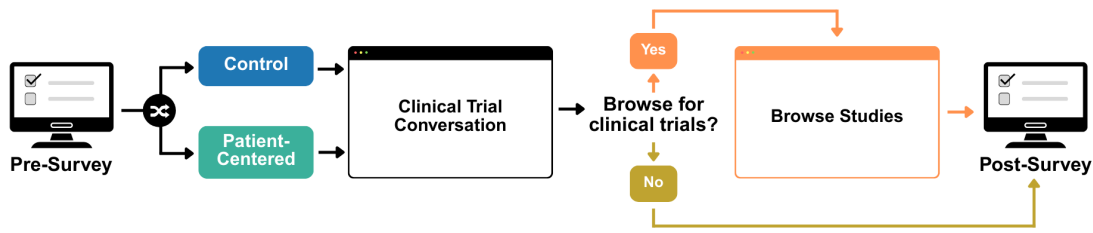


Figure 2: Overview of the study procedure.

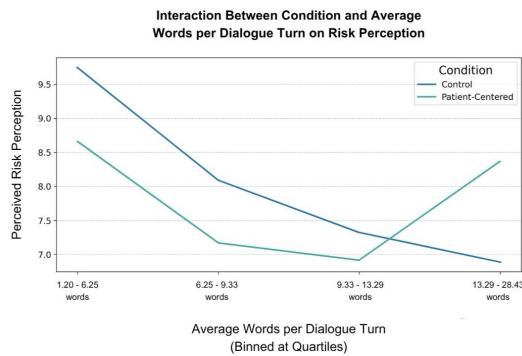


Figure 3: Interaction effect of Condition \times Words on Risk Perception. For visualization purposes, Words per Utterance is divided into four bins based on equal frequencies (quartiles) to illustrate the interaction effect at higher levels of engagement (average words per dialogue turn). Higher values of risk perception indicate less perceived risk.

($M = 12.85, SD = 1.99; EMM = 11.51, SE = 0.25$) reported significantly higher benevolence scores than medium and high literacy participants ($M = 11.45, SD = 2.17; EMM = 12.68, SE = 0.42$), after adjusting for Words per Dialogue Turn. There were no significant interaction effects.

Competence. The final model for Competence did not produce a significant model predicting the relationship between the factors on Competence scores.

| | Model 1 | Model 2 | Model 3 |
|------------------------------------|-----------------|----------------|-----------------|
| (Intercept) | 11.73 (0.31)*** | 12.04 (0.61)** | 11.36 (0.64)*** |
| Condition (Patient-Centered) | 0.19 (0.44) | 0.99 (0.89) | 1.44 (0.98) |
| Words per Dialogue Turn | | -0.03 (0.05) | -0.01 (0.05) |
| Condition \times Words | | -0.08 (0.08) | 0.09 (0.08) |
| Health Literacy (Low) | | | 1.91 (0.67)** |
| Condition \times Health Literacy | | | -1.48 (0.99) |
| R^2 | 0.002 | 0.042 | 0.121 |
| Adj. R^2 | -0.008 | 0.013 | 0.075 |
| F Change | 0.181 | 2.045 | 4.274* |
| F | 0.181 | 1.425 | 2.622* |

Table 3: Multiple regression model for Benevolence with β (standard error). F-statistic with regard to individual model; F Change with regard to prior model step. *** = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.05$

Reciprocity. The final model for Reciprocity was significant (see Table 4 for model steps). Significant main effects emerged only for Health Literacy. Post hoc comparisons using estimated marginal means confirmed that low health literacy participants ($M = 12.70, SD = 2.02; EMM = 12.59, SE = 0.44$) reported significantly higher reciprocity scores than medium and high literacy participants ($M = 11.01, SD = 2.26; EMM = 11.06, SE = 0.26$), after adjusting for Words per Dialogue Turn. There were no significant interaction effects.

| | Model 1 | Model 2 | Model 3 |
|------------------------------------|-----------------|-----------------|-----------------|
| (Intercept) | 11.33 (0.32)*** | 11.67 (0.64)*** | 11.09 (0.67)*** |
| Condition (Patient-Centered) | 0.29 (0.46) | 1.01 (0.94) | 0.87 (1.03) |
| Words per Dialogue Turn | | -0.04 (0.06) | -0.02 (0.05) |
| Condition \times Words | | -0.07 (0.08) | -0.05 (0.08) |
| Health Literacy (Low) | | | 1.67 (0.70)* |
| Condition \times Health Literacy | | | -0.28 (1.04) |
| R^2 | 0.062 | 0.040 | 0.123 |
| Adj. R^2 | -0.006 | 0.010 | 0.077 |
| F Change | 0.381 | 1.823 | 4.513* |
| F | 0.381 | 1.345 | 2.670* |

Table 4: Multiple regression model for Reciprocity with β (standard error). F-statistic with regard to individual model; F Change with regard to prior model step. *** = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.05$

Overall Trust. The final model for Overall Trust as a composite score was significant (see Table 5 for model steps). Significant main effects emerged only for Health Literacy. Post hoc comparisons using estimated marginal means confirmed that low health literacy participants ($M = 49.11, SD = 7.38; EMM = 48.88, SE = 1.18$) reported significantly higher overall trust scores than medium and high literacy participants ($M = 41.89, SD = 5.36; EMM = 42.00, SE = 0.70$), after adjusting for Words per Dialogue Turn. There were no significant interaction effects.

4.2 Browse Behavior.

Browse Choice. The final model for users' decision to browse clinical trials did not produce a significant model predicting the relationship between the factors on Browse Choice.

Follow Up Choice. The final model for Follow Up Choice was significant, ($\chi^2(5) = 13.96, p = 0.016$, Nagelkerke $R^2 = 0.259$) (see Table 6 for model steps). The model correctly classified 71.2% of cases, and The area under the ROC curve was 0.755 (95% CI, 0.628

| | Model 1 | Model 2 | Model 3 |
|------------------------------|-----------------|-----------------|-----------------|
| (Intercept) | 43.29 (0.94)*** | 46.19 (1.84)*** | 43.58 (1.79)*** |
| Condition (Patient-Centered) | 1.10 (1.34) | 0.72 (2.71) | -0.15 (2.75) |
| Words per Dialogue Turn | | -0.29 (0.16) | -0.23 (0.15) |
| Condition × Words | | 0.04 (0.24) | 0.15 (0.22) |
| Health Literacy (Low) | | | 7.28 (1.88)*** |
| Condition × Health Literacy | | | -0.80 (2.78) |
| R^2 | 0.007 | 0.059 | 0.255 |
| Adj. R^2 | -0.003 | 0.030 | 0.216 |
| F Change | 0.669 | 2.697 | 12.531*** |
| F | 0.669 | 2.029 | 6.519*** |

Table 5: Multiple regression model for Overall Trust with β (standard error), F-statistic with regard to individual model; F Change with regard to prior model step. * = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.05$**

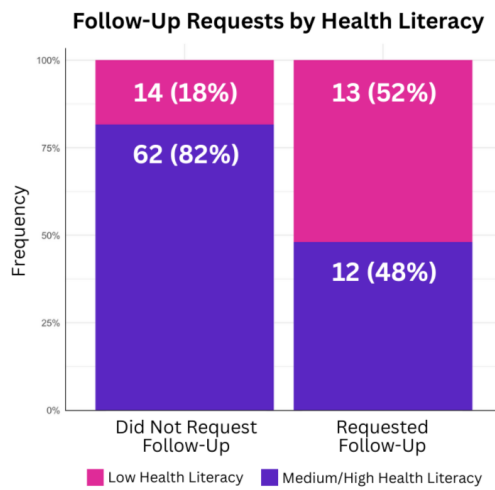


Figure 4: For participants who browsed clinical trials (n = 66), stacked bar graph illustrating the percentage breakdown of participants who did or did not request follow-up

to 0.881), which is an acceptable level of discrimination [31]. Significant main effects emerged only for Health Literacy (see Figure 4). The odds of following up were 18.34 times greater for individuals with low health literacy compared to the medium and high health literacy groups. There were no significant interaction effects.

| | B | SE | Wald | df | p-value | Odds Ratio | 95% CI for Odds Ratio | |
|------------------------------|-------|------|------|----|---------|------------|-----------------------|--------|
| | | | | | | | Lower | Upper |
| (Intercept) | -0.78 | 0.91 | 0.74 | 1 | 0.391 | 0.46 | | |
| Condition (Patient-Centered) | 0.50 | 1.27 | 0.16 | 1 | 0.694 | 1.65 | 0.14 | 19.85 |
| Words per Dialogue Turn | -0.07 | 0.08 | 0.77 | 1 | 0.382 | 0.93 | 0.80 | 1.09 |
| Condition × Words | 0.02 | 0.11 | 0.03 | 1 | 0.864 | 1.02 | 0.82 | 1.26 |
| Health Literacy (Low) | 2.91 | 1.01 | 8.37 | 1 | 0.004** | 18.34 | 2.56 | 131.65 |
| Condition × Health Literacy | -1.98 | 1.29 | 2.35 | 1 | 0.125 | 0.14 | 0.01 | 1.74 |

Table 6: Binomial logistic regression model for Follow Up Choice for participants who browsed clinical trials (n = 66)

5 Discussion

This study explored the design of a virtual agent intervention for engaging breast cancer patients with clinical trials. Specifically,

we investigated how implementing patient-centered communication strategies in a virtual agent influenced perceived trust (RQ1) and users' browse behaviors (RQ2). Our findings reveal that the patient-centered virtual agent alone did not improve trust, nor did it influence users' browse behavior. However, the inclusion of users' health literacy (RQ1a) and engagement (RQ2a) indicated a more nuanced relationship with the perception of trust and browse behavior. We discuss these findings below.

5.1 The Impact of Health Literacy

Our findings revealed that low health literacy predicted improved overall trust, specifically for risk perception, benevolence, and reciprocity. Prior research suggests that people with low health literacy tend to trust information from television, social media, and blogs more than from institutional or medical websites, which are preferred by those with higher literacy [14]. This preference for conversational and accessible formats may explain why low health literacy participants in our study perceived the virtual agent – visually dynamic and interactive – as less risky and more benevolent than their higher literacy counterparts. Additionally, prior work suggests that low health literacy users find animated virtual agents to be engaging, preferring them over traditional information delivery methods [7, 8]. The greater reciprocity reported by low literacy participants in our study suggests they may have experienced a stronger social connection with the agent. This aligns with the Computers as Social Actors framework [49] and underscores the relational potential of virtual agents for enhancing patient education among low health literacy users.

Health literacy also significantly predicted the browse behaviors of participants. Specifically, of those who browsed (n=66), low health literacy participants were more likely to request follow-up information compared to medium and high health literacy participants. This may have been influenced by increased trust low health literacy participants had in the virtual agent, leading them to seek more information on what the virtual agent shared after the interaction. Another possibility is that low health literacy participants may have needed additional information to better understand what they learned during the intervention, as prior work suggests that access to more information helps lower health literacy populations understand medical and scientific information [26, 36, 40, 65]. Although our study does not definitively explain why low health literacy participants browsed more, these findings suggest that the virtual agent intervention was particularly effective in encouraging this group to seek additional information about clinical trials compared to participants with medium and high health literacy.

5.2 The Impact of a Patient-Centered Virtual Agent

Interestingly, our findings revealed that the patient-centered condition alone had an overall negative impact on risk perception. However, our analysis also revealed a more complex interaction between agent condition and user engagement.

In the control condition, users with high engagement (13 to 28 words per dialogue turn) tended to perceive the agent as increasingly risky. In contrast, in the patient-centered condition, users with similar length responses reported lower perceived risk. When users

813 contributed more elaborate responses, the agent may have had
814 more information to tailor its linguistic style and emotional expres-
815 sion, as the LIWC validation showed higher emotional and stylistic
816 alignment in the patient-centered condition. However, users who
817 engaged at lower levels may not have provided enough input to
818 trigger the agent’s adaptive mechanisms, leading to less effective
819 personalization and, in turn, higher perceived risk.

820 The effectiveness of the patient-centered agent at high levels
821 of user engagement mirrors prior work suggesting that adaptive
822 systems are most effective when users are emotionally invested or
823 highly engaged [56]. Users who typed longer responses possibly
824 offered more emotional and contextual cues, enabling the agent
825 to better personalize its replies through empathetic language and
826 linguistic style matching, as our LIWC validation showed higher
827 emotional and stylistic alignment in the patient-centered condition.
828 Prior work demonstrates that when a virtual agent aligns its com-
829 munication to the patient, it is perceived as more trustworthy [92],
830 consistent with mimicry and attunement effects known to enhance
831 interpersonal trust [??]. In our study, emotional language in par-
832 ticular may have helped convey concern and affective presence,
833 making the agent appear more trustworthy. In sum, our results sug-
834 gest that the patient-centered virtual agent’s impact on reducing
835 perceived risk may rely on user engagement.

837 5.3 Design Guidelines and Takeaways

838 Our findings provide several actionable insights for designing vir-
839 tual agent interventions. Although the context of our work is breast
840 cancer patients and clinical trials, our focus on communicating
841 complex health information makes our insights valuable for appli-
842 cations in health education.

843
844
845 *Design with Health Literacy in Mind.* Health literacy emerged
846 as a key moderator in our study, with participants reporting low
847 health literacy perceiving the virtual agent as more trustworthy and
848 showing greater interest in follow-up information. This suggests
849 that virtual agents may be especially effective for populations that
850 often face barriers in accessing and trusting complex health infor-
851 mation and groups that are disproportionately underrepresented
852 in clinical trials [93]. Additionally, prior work highlights a need for
853 online interventions to connect this population to credible informa-
854 tion sources [36] as language barriers can lead lower health literacy
855 users to rely on misinformation and even reduce engagement with
856 proper healthcare resources [19]. To this end, our work shows that
857 virtual agent interventions have the potential to be designed not
858 only to present more readable content but also seamlessly link users
859 to high-quality resources. Therefore, healthcare interventions can
860 leverage virtual agents to improve inclusion for low health literacy
861 users and streamline access to credible information, empowering
862 them to engage with high-quality health resources.

863
864 *Adapt to User Engagement Levels.* Our findings revealed that
865 the effectiveness of patient-centered virtual agents in reducing per-
866 ceived risk depends heavily on the user’s level of engagement. Users
867 who provided longer input reported lower perceived risk, as the
868 agent may have been better able to align emotionally and linguisti-
869 cally [56]. Designers could therefore account for varying user input

870 by creating agents that can detect and adapt to different engage-
871 ment levels in real time with strategies such as asking follow-up
872 questions and encouraging elaboration [52]. By dynamically tailor-
873 ing patient-centered behavior to user engagement levels, virtual
874 agents can maintain alignment and empathy across diverse inter-
875 action styles. This could ensure that adaptive efforts work across
876 engagement levels, ultimately helping to further reduce perceived
877 risk.

880 6 Limitations and Future Work

881 This study provides valuable insights into the role of health lit-
882 eracy and patient-centered communication strategies in virtual
883 agent interventions for breast cancer clinical trial engagement, but
884 several limitations remain. First, our participants were recruited
885 online, limiting generalizability to populations with limited internet
886 access or technological proficiency. Future work should evaluate
887 virtual agents in more diverse settings to ensure equitable applica-
888 bility across diverse patient populations. Second, while we focused
889 on health literacy due to its relevance in cancer communication,
890 other user characteristics – such as race, ethnicity, age, personal-
891 ity, and prior healthcare experiences – may also impact outcomes.
892 Our participants were predominantly White and female, limiting
893 generalizability to other health applications and contexts. Future
894 studies should examine other contexts and a broader range of user
895 traits. Third, our work focused on a single intervention type with-
896 out direct comparisons to other approaches, such as text-based
897 interventions. Future work should explore the relative strengths
898 of virtual agents for low health literacy populations. Finally, this
899 study did not investigate outcomes such as intention to enroll in
900 clinical trials. While our work explored important precursors such
901 as trust and intervention behaviors, future work should examine
902 longer-term effects to evaluate real-world impact.

903 In conclusion, while this study highlights the promise of virtual
904 agents in addressing barriers to clinical trial participation, further
905 research is needed to refine their design and evaluate their practical
906 use. Addressing these limitations enables future efforts to maximize
907 the potential of virtual agent health-related interventions.

910 7 Conclusion

911 This study explored patient-centered communication in virtual
912 agents and the impact of users’ health literacy in the design of
913 virtual agent interventions aimed at improving clinical trial en-
914 gagement among breast cancer patients. Our findings suggest that
915 patient-centered communication strategies alone may not enhance
916 trust or drive intervention outcomes. Instead, health literacy strongly
917 predicted trust and engagement, revealing that the effectiveness of
918 communication strategies is moderated by both literacy and user
919 engagement. Our work highlights the potential of virtual agents
920 to address critical gaps for those with low health literacy. By pri-
921 oritizing the integration of user characteristics like health literacy,
922 virtual agents can better support patients in navigating complex
923 healthcare topics, such as clinical trial participation. These interven-
924 tions hold promise not only for improving trust and engagement
925 but also for promoting equity for advancements in cancer research
926 and treatment.

References

- [1] Amal Abdulrahman, Deborah Richards, and Ayse Aysin Bilgin. 2022. Exploring the influence of a user-specific explainable virtual advisor on health behaviour change intentions. *Autonomous Agents and Multi-Agent Systems* 36, 1 (2022), 25.
- [2] Reza Amini, Christine Lisetti, and Ugan Yasavur. 2014. Emotionally responsive virtual counselor for behavior-change health interventions. In *Advancing the Impact of Design Science: Moving from Theory to Practice: 9th International Conference, DESRIST 2014, Miami, FL, USA, May 22-24, 2014. Proceedings 9*. Springer, 433–437.
- [3] Simon Arvidsson and Johan Axell. 2023. Prompt engineering guidelines for LLMs in Requirements Engineering. (2023).
- [4] Nancy E Avis, Kevin W Smith, Carol L Link, Gabriel N Hortobagyi, and Edgardo Rivera. 2006. Factors associated with participation in breast cancer treatment clinical trials. *Journal of Clinical Oncology* 24, 12 (2006), 1860–1867.
- [5] Jose Barambones, Cristian Moral, Angélica de Antonio, Ricardo Imbert, Loïc Martínez, and Elena Villalba-Mora. 2024. ChatGPT for learning HCI techniques: A Case Study on Interviews for Personas. *IEEE Transactions on Learning Technologies* (2024).
- [6] Federica Biassoni and Martina Gnerre. 2025. Exploring ChatGPT’s communication behaviour in healthcare interactions: A psycholinguistic perspective. *Patient Education and Counseling* 134 (2025), 108663.
- [7] Timothy W Bickmore, Laura M Pfeifer, and Brian W Jack. 2009. Taking the time to care: empowering low health literacy hospital patients with virtual nurse agents. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1265–1274.
- [8] Timothy W Bickmore, Laura M Pfeifer, and Michael K Paasche-Orlow. 2009. Using computer agents to explain medical documents to patients with low health literacy. *Patient education and counseling* 75, 3 (2009), 315–320.
- [9] Timothy W Bickmore, Daniel Schulman, and Candace L Sidner. 2011. A reusable framework for health counseling dialogue systems based on a behavioral medicine ontology. *Journal of biomedical informatics* 44, 2 (2011), 183–197.
- [10] Timothy W Bickmore, Dina Utami, Robin Matsuyama, and Michael K Paasche-Orlow. 2016. Improving access to online health information with conversational agents: a randomized controlled experiment. *Journal of medical Internet research* 18, 1 (2016), e1.
- [11] Emma G Bryan, Huan Chen, Melissa Vilaro, Haoran Chu, Gabriella Grillo, Palani Te, Miriam Buhr, Stephen Anton, and Janice L Krieger. 2025. Developing a supportive virtual human to deliver clinical trial education for older women and other populations historically excluded from research. *Patient Education and Counseling* 130 (2025), 108485.
- [12] Christopher Burton, Aurora Szentagotai Tatar, Brian McKinstry, Colin Matheson, Silviu Matu, Ramona Moldovan, Michele Macnab, Elaine Farrow, Daniel David, Claudia Pagliari, et al. 2016. Pilot randomised controlled trial of Help4Mood, an embodied virtual agent-based system to support treatment of depression. *Journal of telemedicine and telecare* 22, 6 (2016), 348–355.
- [13] Justine Cassell and Timothy Bickmore. 2000. External manifestations of trustworthiness in the interface. *Commun. ACM* 43, 12 (2000), 50–56.
- [14] Xuewei Chen, Jennifer L Hay, Erika A Waters, Marc T Kiviniemi, Caitlin Biddle, Elizabeth Schofield, Yuelin Li, Kimberly Kaphingst, and Heather Orom. 2018. Health literacy and use and trust in health information. *Journal of health communication* 23, 8 (2018), 724–734.
- [15] ClinicalTrials.gov. 2025. Study NCT04990895: A Clinical Study Available on ClinicalTrials.gov. <https://clinicaltrials.gov/study/NCT04990895>. Accessed: January 20, 2025.
- [16] ClinicalTrials.gov. 2025. Study NCT05488145: A Clinical Study Available on ClinicalTrials.gov. <https://clinicaltrials.gov/study/NCT05488145>. Accessed: January 20, 2025.
- [17] ClinicalTrials.gov. 2025. Study NCT05868187: A Clinical Study Available on ClinicalTrials.gov. <https://clinicaltrials.gov/study/NCT05868187>. Accessed: January 20, 2025.
- [18] Eric J Cooks, Kyle A Duke, Elizabeth Flood-Grady, Melissa J Vilaro, Rashi Ghosh, Naomi Parker, Palani Te, Thomas J George, Benjamin C Lok, Maribeth Williams, et al. 2022. Can virtual human clinicians help close the gap in colorectal cancer screening for rural adults in the United States? The influence of rural identity on perceptions of virtual human clinicians. *Preventive Medicine Reports* 30 (2022), 102034.
- [19] Nicola Diviani, Bas Van Den Putte, Stefano Giani, and Julia CM van Weert. 2015. Low health literacy and evaluation of online health information: a systematic review of the literature. *Journal of medical Internet research* 17, 5 (2015), e112.
- [20] Lucile Dupuy, Jean-Arthur Micoulaud-Franchi, and Pierre Philip. 2021. Acceptance of virtual agents in a homecare context: Evaluation of excessive daytime sleepiness in apneic patients during interventions by continuous positive airway pressure (CPAP) providers. *Journal of sleep research* 30, 2 (2021), e13094.
- [21] Joy Egede, Maria J Galvez Trigo, Adrian Hazzard, Martin Porcheron, Edgar Bodiaj, Joel E Fischer, Chris Greenhalgh, and Michel Valstar. 2021. Designing an adaptive embodied conversational agent for health literacy: a user study. In *Proceedings of the 21st ACM International Conference on Intelligent Virtual Agents*. 112–119.
- [22] Kiameesha R Evans, M Jane Lewis, and Shawna V Hudson. 2012. The role of health literacy on African American and Hispanic/Latino perspectives on cancer clinical trials. *Journal of Cancer Education* 27 (2012), 299–305.
- [23] Angelica Falkenstein, Brandon Tran, Daniel Ludi, Afshin Molkara, Henry Nguyen, Arnold Tabuenca, and Kate Sweeny. 2016. Characteristics and correlates of word use in physician-patient communication. *Annals of Behavioral Medicine* 50, 5 (2016), 664–677.
- [24] Daniela B Friedman, Sei-Hill Kim, Andrea Tanner, Caroline D Bergeron, Caroline Foster, and Kevin General. 2014. How are we communicating about clinical trials?: an assessment of the content and readability of recruitment resources. *Contemporary clinical trials* 38, 2 (2014), 275–283.
- [25] Paula M Gardiner, Kelly D McCue, Lily M Negash, Teresa Cheng, Laura F White, Leanne Yinusa-Nyahkoon, Brian W Jack, and Timothy W Bickmore. 2017. Engaging women with an embodied conversational agent to deliver mindfulness and lifestyle recommendations: A feasibility randomized control trial. *Patient education and counseling* 100, 9 (2017), 1720–1729.
- [26] Patrick R Gradie, Megan Litster, Rinu Thomas, Jay Vyas, and Martin R Schiller. 2011. SciReader enables reading of medical content with instantaneous definitions. *BMC medical informatics and decision making* 11 (2011), 1–5.
- [27] Siddharth Gulati, Sonia Sousa, and David Lamas. 2019. Design, development and evaluation of a human-computer trust scale. *Behaviour & Information Technology* 38, 10 (2019), 1004–1015.
- [28] Emma Hagens, Maria Luce Lupetti, Willem van der Maden, Régine Steegers, and Melek Rousian. 2023. Trustworthy Embodied Conversational Agents for Healthcare. In *5th Conference on Conversational User Interfaces*. <https://doi.org/10.1145/3571884.3597128>.
- [29] Jolie Haun, Virginia Noland-Dodd, Jill Varnes, John Graham-Pole, Barbara Rienzo, and Patricia Donaldson. 2009. Testing the BRIEF health literacy screening tool. *Fed Pract* 26, 12 (2009), 24–31.
- [30] Hyehyun Hong and Hyun Jee Oh. 2020. The effects of patient-centered communication: exploring the mediating role of trust in healthcare providers. *Health communication* 35, 4 (2020), 502–511.
- [31] David W Hosmer Jr, Stanley Lemeshow, and Rodney X Sturdivant. 2013. *Applied logistic regression*. John Wiley & Sons.
- [32] Robert H Houlihan, Michael H Kennedy, Robert R Kulesher, Stephenie C Lemon, D Lawrence Wickerham, Chung-Cheng Hsieh, and Dario C Altieri. 2010. Identification of accrual barriers onto breast cancer prevention clinical trials: A case-control study. *Cancer* 116, 15 (2010), 3569–3576.
- [33] Shaohai Jiang, Zhengyu Wu, Xiaoyu Zhang, Ying Ji, Jingyong Xu, Peng Liu, Yan Liu, Jie Zheng, Liang Zhao, and Jingxi Chen. 2024. How does patient-centered communication influence patient trust?: The roles of patient participation and patient preference. *Patient Education and Counseling* 122 (2024), 108161.
- [34] Emma Kemp, Joshua Trigg, Lisa Beatty, Chris Christensen, Haryana M Dhillon, Anthony Maeder, Patricia AH Williams, and Bogda Koczwara. 2021. Health literacy, digital health literacy and the implementation of digital health technologies in cancer care: the need for a strategic approach. *Health Promotion Journal of Australia* 32 (2021), 104–114.
- [35] B Price Kerfoot, Michael C Kearney, Donna Connelly, and Michael L Ritchey. 2009. Interactive spaced education to assess and improve knowledge of clinical practice guidelines: a randomized controlled trial.
- [36] Henna Kim and Bo Xie. 2017. Health literacy in the eHealth era: a systematic review of the literature. *Patient education and counseling* 100, 6 (2017), 1073–1082.
- [37] Ann King and Ruth B Hoppe. 2013. “Best practice” for patient-centered communication: a narrative review. *Journal of graduate medical education* 5, 3 (2013), 385–393.
- [38] Liliانا Laranjo, Adam G Dunn, Huong Ly Tong, Ahmet Baki Kocaballi, Jessica Chen, Rabia Bashir, Didi Surian, Blanca Gallego, Farah Magrabi, Annie YS Lau, et al. 2018. Conversational agents in healthcare: a systematic review. *Journal of the American Medical Informatics Association* 25, 9 (2018), 1248–1258.
- [39] Guek Eng Lee, Mandy Ow, Desiree Lie, and Rebecca Dent. 2016. Barriers and facilitators for clinical trial participation among diverse Asian patients with breast cancer: a qualitative study. *BMC women’s health* 16 (2016), 1–8.
- [40] Gonyd Leroy and Trudi Miller. 2010. Perils of providing visual health information overviews for consumers with low health literacy or high stress. *Journal of the American Medical Informatics Association* 17, 2 (2010), 220–223.
- [41] Wendy Levinson, Cara S Lesser, and Ronald M Epstein. 2010. Developing physician communication skills for patient-centered care. *Health affairs* 29, 7 (2010), 1310–1318.
- [42] LIWC. 2025. LIWC: Language Style Matching. <https://www.liwc.app/help/lsm/>. Accessed: January 20, 2025.
- [43] LIWC. 2025. LIWC: Linguistic Inquiry and Word Count. <https://www.liwc.app/>. Accessed: January 20, 2025.
- [44] Julie B Mallinger, Jennifer J Griggs, and Cleveland G Shields. 2005. Patient-centered care and breast cancer survivors’ satisfaction with information. *Patient education and counseling* 57, 3 (2005), 342–349.
- [45] Julien Mancini, Jean Genève, Florence Dalenc, Dominique Genre, Alain Monnier, Pierre Kerbrat, Rémy Largillier, Daniel Serin, Maria Rios, Henri Roché, et al. 2007. Decision-making and breast cancer clinical trials: how experience challenges

- attitudes. *Contemporary clinical trials* 28, 6 (2007), 684–694.
- [46] Andrew Maxim, Mohan Zalake, and Benjamin Lok. 2023. The Impact of Virtual Human Vocal Personality on Establishing Rapport: A Study on Promoting Mental Wellness Through Extroversion and Vocalics. In *Proceedings of the 23rd ACM International Conference on Intelligent Virtual Agents*. 1–8.
- [47] Kathleen M Mazor, René L Beard, Gwen L Alexander, Neeraj K Arora, Cassandra Firreno, Bridget Gaglio, Sarah M Greene, Celeste A Lemay, Brandi E Robinson, Douglas W Roblin, et al. 2013. Patients' and family members' views on patient-centered communication during cancer care. *Psycho-Oncology* 22, 11 (2013), 2487–2495.
- [48] Sharon Mozgai, Arno Hartholt, Dayo Akinyemi, Katarina Kubicek, Albert Rizzo, and Michele Kipke. 2020. Development and Initial Feasibility Testing of the Virtual Research Navigator (VRN): A Public-Facing Agent-Based Educational System for Clinical Research Participation. In *HCI International 2020-Posters: 22nd International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings, Part II 22*. Springer, 304–307.
- [49] Clifford Nass, Jonathan Steuer, and Ellen R Tauber. 1994. Computers are social actors. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 72–78.
- [50] National Cancer Institute. 2025. Cancer Stat Facts: Common Cancer Sites. <https://seer.cancer.gov/staffacts/html/common.html>. Accessed: January 18, 2025.
- [51] Mark Neerincx and Jan Willem Streefkerk. 2003. Interacting in desktop and mobile context: Emotion, trust, and task performance. In *European Symposium on Ambient Intelligence*. Springer, 119–132.
- [52] Catharine Oertel, Ginevra Castellano, Mohamed Chetouani, Jauwairia Nasir, Mohammad Obaid, Catherine Pelachaud, and Christopher Peters. 2020. Engagement in human-agent interaction: An overview. *Frontiers in Robotics and AI* 7 (2020), 92.
- [53] Stefan Olafsson, Teresa K O'Leary, and Timothy W Bickmore. 2020. Motivating health behavior change with humorous virtual agents. In *Proceedings of the 20th ACM international conference on intelligent virtual agents*. 1–8.
- [54] OpenAI. 2025. OpenAI Platform: Assistants Overview. <https://platform.openai.com/docs/assistants/overview>. Accessed: January 20, 2025.
- [55] Serena Petrocchi, Paola Iannello, Flavia Lecciso, Annalisa Levante, Alessandro Antonietti, and PJ Schulz. 2019. Interpersonal trust in doctor-patient relation: Evidence from dyadic analysis and association with quality of dyadic communication. *Social science & medicine* 235 (2019), 112391.
- [56] Hedieh Ranjbarbar, Deborah Richards, Ayse Aysin Bilgin, and Cat Kutay. 2019. First impressions count! The role of the human's emotional state on rapport established with an empathic versus neutral virtual therapist. *IEEE transactions on affective computing* 12, 3 (2019), 788–800.
- [57] Janet D Record, Roy C Ziegelstein, Colleen Christmas, Cynthia S Rand, and Laura A Hanyok. 2021. Delivering personalized care at a distance: how telemedicine can foster getting to know the patient as a person. *Journal of Personalized Medicine* 11, 2 (2021), 137.
- [58] Byron Reeves and Clifford Nass. 1996. The media equation: How people treat computers, television, and new media like real people. *Cambridge, UK* 10, 10 (1996), 19–36.
- [59] Minjin Rheu, Ji Youn Shin, Wei Peng, and Jina Huh-Yoo. 2021. Systematic review: Trust-building factors and implications for conversational agent design. *International Journal of Human-Computer Interaction* 37, 1 (2021), 81–96.
- [60] Daniel Schulman and Timothy Bickmore. 2009. Persuading users through counseling dialogue with a conversational agent. In *Proceedings of the 4th international conference on persuasive technology*. 1–8.
- [61] Timothy Shaw, Andrea Long, Sanjiv Chopra, and B Price Kerfoot. 2011. Impact on clinical behavior of face-to-face continuing medical education blended with online spaced education: a randomized controlled trial. *Journal of Continuing Education in the Health Professions* 31, 2 (2011), 103–108.
- [62] Rebecca L Siegel, Angela N Giaquinto, and Ahmedin Jemal. 2024. Cancer statistics, 2024. *CA: a cancer journal for clinicians* 74, 1 (2024), 12–49.
- [63] Jeannine S Skinner, Alecia M Fair, Alexis S Holman, Alaina P Boyer, and Consuelo H Wilkins. 2019. The impact of an educational video on clinical trial enrollment and knowledge in ethnic minorities: a randomized control trial. *Frontiers in Public Health* 7 (2019), 104.
- [64] Rita Sommers, Cornelius Van Staden, and Francois Steffens. 2017. Views of clinical trial participants on the readability and their understanding of informed consent documents. *AJOB Empirical Bioethics* 8, 4 (2017), 277–284.
- [65] Colin M Sox, William M Gribbons, Beth A Loring, Kenneth D Mandl, Rosanna Batista, and Stephen C Porter. 2010. Patient-centered design of an information management module for a personally controlled health record. *Journal of Medical Internet Research* 12, 3 (2010), e1269.
- [66] Ian Steenstra, Prasanth Murali, Rebecca Perkins, Natalie Joseph, Michael Paasche-Orlow, and Timothy Bickmore. 2023. Changing Parent Attitudes Towards HPV Vaccination by Including Adolescents in Multiparty Counseling using Virtual Agents. In *Proceedings of the 23rd ACM International Conference on Intelligent Virtual Agents*. 1–8.
- [67] Ian Steenstra, Farnaz Nouraei, Mehdi Arjmand, and Timothy Bickmore. 2024. Virtual agents for alcohol use counseling: Exploring llm-powered motivational interviewing. In *Proceedings of the ACM International Conference on Intelligent Virtual Agents*. 1–10.
- [68] Mary M Step, Julia Hannum Rose, Jeffrey M Albert, Vinay K Cheruvu, and Laura A Siminoff. 2009. Modeling patient-centered communication: oncologist relational communication and patient communication involvement in breast cancer adjuvant therapy decision-making. *Patient education and counseling* 77, 3 (2009), 369–378.
- [69] Carolin Straßmann, Nicole C Krämer, Hendrik Buschmeier, and Stefan Kopp. 2020. Age-related differences in the evaluation of a virtual health agent's appearance and embodiment in a health-related interaction: experimental lab study. *Journal of medical Internet research* 22, 4 (2020), e13726.
- [70] Richard L Street Jr, Howard S Gordon, Michael M Ward, Edward Krupat, and Richard L Kravitz. 2005. Patient participation in medical consultations: why some patients are more involved than others. *Medical care* 43, 10 (2005), 960–969.
- [71] Guangzhi Sun, Xiao Zhan, and Jose Such. 2024. Building better ai agents: A provocation on the utilisation of persona in llm-based conversational agents. In *Proceedings of the 6th ACM Conference on Conversational User Interfaces*. 1–6.
- [72] Synthesia. 2025. Synthesia: AI Video Creation Platform. <https://www.synthesia.io/>. Accessed: January 20, 2025.
- [73] Silke Ter Stal, Gerbrich Jongbloed, and Monique Tabak. 2021. Embodied conversational agents in eHealth: How facial and textual expressions of positive and neutral emotions influence perceptions of mutual understanding. *Interacting with Computers* 33, 2 (2021), 167–176.
- [74] Silke ter Stal, Monique Tabak, Harm op den Akker, Tessa Beinema, and Hermie Hermens. 2020. Who do you prefer? The effect of age, gender and role on users' first impressions of embodied conversational agents in eHealth. *International Journal of Human-Computer Interaction* 36, 9 (2020), 881–892.
- [75] Amelia A Trant, Lucas Walz, Whitney Allen, Jose DeJesus, Christos Hatzis, and Andrea Silber. 2020. Increasing accrual of minority patients in breast cancer clinical trials. *Breast Cancer Research and Treatment* 184 (2020), 499–505.
- [76] Joseph M Unger, Dawn L Hershman, Cathee Till, Lori M Minasian, Raymond U Osarogiagbon, Mark E Fleury, and Riha Vaidya. 2021. "When offered to participate": a systematic review and meta-analysis of patient agreement to participate in cancer clinical trials. *JNCI: Journal of the National Cancer Institute* 113, 3 (2021), 244–257.
- [77] Joseph M Unger, Lawrence N Shulman, Matthew A Facktor, Heidi Nelson, and Mark E Fleury. 2024. National estimates of the participation of patients with cancer in clinical research studies based on Commission on cancer accreditation data. *Journal of Clinical Oncology* (2024), JCO-23.
- [78] Martien JP van Bussel, Gaby J Odekerken-Schröder, Carol Ou, Rachelle R Swart, and Maria JG Jacobs. 2022. Analyzing the determinants to accept a virtual assistant and use cases among cancer patients: a mixed methods study. *BMC Health Services Research* 22, 1 (2022), 890.
- [79] Maria K Venetis, Jeffrey D Robinson, Katie LaPlant Turkiewicz, and Mike Allen. 2009. An evidence base for patient-centered cancer care: a meta-analysis of studies of observed communication between cancer specialists and their patients. *Patient education and counseling* 77, 3 (2009), 379–383.
- [80] Melissa J Vilaro, Danyell S Wilson-Howard, Lauren N Griffin, Fatemeh Tavassoli, Mohan S Zalake, Benjamin C Lok, Francois P Modave, Thomas J George, Peter J Carek, and Janice L Krieger. 2020. Tailoring virtual human-delivered interventions: A digital intervention promoting colorectal cancer screening for Black women. *Psycho-Oncology* 29, 12 (2020), 2048–2056.
- [81] Melissa J Vilaro, Danyell S Wilson-Howard, Mohan S Zalake, Fatemeh Tavassoli, Benjamin C Lok, Francois P Modave, Thomas J George, Folakemi Odedina, Peter J Carek, and Janice L Krieger. 2021. Key changes to improve social presence of a virtual health assistant promoting colorectal cancer screening informed by a technology acceptance model. *BMC medical informatics and decision making* 21, 1 (2021), 196.
- [82] Catharine Wang, Timothy Bickmore, Deborah J Bowen, Tricia Norkunas, MaryAnn Campion, Howard Cabral, Michael Winter, and Michael Paasche-Orlow. 2015. Acceptability and feasibility of a virtual counselor (VICKY) to collect family health histories. *Genetics in Medicine* 17, 10 (2015), 822–830.
- [83] Siqi Wang, Xiaofei Zhang, and Fanbo Meng. 2023. What leads to effective online physician-patient communication? The power of convergence. In *Wuhan International Conference on E-business*. Springer, 331–343.
- [84] Anna Winkler, Patrick Kutschar, Stefan Pitzer, Antje van der Zee-Neuen, Susanne Kerner, Jürgen Osterbrink, and Simon Krutter. 2023. Avatar and virtual agent-assisted telecare for patients in their homes: A scoping review. *Journal of Telemedicine and Telecare* (2023), 1357633X231174484.
- [85] Charlene A Wong, William B Song, Megan Jiao, Emily O'Brien, Peter Ubel, Gary Wang, and Charles D Scales. 2021. Strategies for research participant engagement: a synthetic review and conceptual framework. *Clinical Trials* 18, 4 (2021), 457–465.
- [86] Christopher You, Rashi Ghosh, Andrew Maxim, Jacob Stuart, Eric Cooks, and Benjamin Lok. 2022. How does a virtual human earn your trust? guidelines to improve willingness to self-disclose to intelligent virtual agents. In *Proceedings of the 22nd ACM International Conference on Intelligent Virtual Agents*. 1–8.

| | | | |
|------|------|--|------|
| 1161 | [87] | Mohan Zalake, Alexandre Gomes De Siqueira, Krishna Vaddiparti, Pavlo Antonenko, and Benjamin Lok. 2021. Towards understanding how virtual human’s verbal persuasion strategies influence user intentions to perform health behavior. In <i>Proceedings of the 21st ACM International Conference on Intelligent Virtual Agents</i> . 216–223. | 1219 |
| 1162 | | | 1220 |
| 1163 | | | 1221 |
| 1164 | | | 1222 |
| 1165 | [88] | Mohan Zalake, Alexandre Gomes De Siqueira, Krishna Vaddiparti, Pavlo Antonenko, and Benjamin Lok. 2024. Evaluating the Effect of Adapting Virtual Humans Based on Individual Differences in Users. In <i>International Conference on Human-Computer Interaction</i> . Springer, 405–423. | 1223 |
| 1166 | | | 1224 |
| 1167 | [89] | Mohan Zalake, Fatemeh Tavassoli, Lauren Griffin, Janice Krieger, and Benjamin Lok. 2019. Internet-based tailored virtual human health intervention to promote colorectal cancer screening: design guidelines from two user studies. In <i>Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents</i> . 73–80. | 1225 |
| 1168 | | | 1226 |
| 1169 | | | 1227 |
| 1170 | | | 1228 |
| 1171 | [90] | Lingling Zhang, Sabarish V Babu, Meenu Jindal, Joel E Williams, Ronald W Gimbel, et al. 2019. A patient-centered mobile phone app (iHeartU) with a virtual human assistant for self-management of heart failure: protocol for a usability assessment study. <i>JMIR research protocols</i> 8, 5 (2019), e13502. | 1229 |
| 1172 | | | 1230 |
| 1173 | | | 1231 |
| 1174 | [91] | Zhe Zhang and Timothy Bickmore. 2018. Medical shared decision making with a virtual agent. In <i>Proceedings of the 18th International Conference on Intelligent Virtual Agents</i> . 113–118. | 1232 |
| 1175 | | | 1233 |
| 1176 | | | 1234 |
| 1177 | | | 1235 |
| 1178 | | | 1236 |
| 1179 | | | 1237 |
| 1180 | | | 1238 |
| 1181 | | | 1239 |
| 1182 | | | 1240 |
| 1183 | | | 1241 |
| 1184 | | | 1242 |
| 1185 | | | 1243 |
| 1186 | | | 1244 |
| 1187 | | | 1245 |
| 1188 | | | 1246 |
| 1189 | | | 1247 |
| 1190 | | | 1248 |
| 1191 | | | 1249 |
| 1192 | | | 1250 |
| 1193 | | | 1251 |
| 1194 | | | 1252 |
| 1195 | | | 1253 |
| 1196 | | | 1254 |
| 1197 | | | 1255 |
| 1198 | | | 1256 |
| 1199 | | | 1257 |
| 1200 | | | 1258 |
| 1201 | | | 1259 |
| 1202 | | | 1260 |
| 1203 | | | 1261 |
| 1204 | | | 1262 |
| 1205 | | | 1263 |
| 1206 | | | 1264 |
| 1207 | | | 1265 |
| 1208 | | | 1266 |
| 1209 | | | 1267 |
| 1210 | | | 1268 |
| 1211 | | | 1269 |
| 1212 | | | 1270 |
| 1213 | | | 1271 |
| 1214 | | | 1272 |
| 1215 | | | 1273 |
| 1216 | | | 1274 |
| 1217 | | | 1275 |
| 1218 | | | 1276 |

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009