Review

Computer-Controlled Virtual Humans in Patient-Facing Systems: Systematic Review and Meta-Analysis

Debaleena Chattopadhyay[1], PhD; Tengteng Ma[2], MSc; Hasti Sharifi[1], BSc; Pamela Martyn-Nemeth[3], PhD

[1]Department of Computer Science, University of Illinois at Chicago, Chicago, IL, United States

[2]Department of Information and Decision Sciences, University of Illinois at Chicago, Chicago, IL, United States

[3]Department of Biobehavioral Health Science, University of Illinois at Chicago, Chicago, IL, United States

Corresponding Author:
Debaleena Chattopadhyay, PhD
Department of Computer Science
University of Illinois at Chicago
851 S Morgan (M/C 152)
Room 1120 SEO
Chicago, IL, 60607
United States
Phone: 1 312413518
Email: debchatt@uic.edu

Abstract

Background: Virtual humans (VH) are computer-generated characters that appear humanlike and simulate face-to-face conversations using verbal and nonverbal cues. Unlike formless conversational agents, like smart speakers or chatbots, VH bring together the capabilities of both a conversational agent and an interactive avatar (computer-represented digital characters). Although their use in patient-facing systems has garnered substantial interest, it is unknown to what extent VH are effective in health applications.

Objective: The purpose of this review was to examine the effectiveness of VH in patient-facing systems. The design and implementation characteristics of these systems were also examined.

Methods: Electronic bibliographic databases were searched for peer-reviewed articles with relevant key terms. Studies were included in the systematic review if they designed or evaluated VH in patient-facing systems. Of the included studies, studies that used a randomized controlled trial to evaluate VH were included in the meta-analysis; they were then summarized using the PICOTS framework (population, intervention, comparison group, outcomes, time frame, setting). Summary effect sizes, using random-effects models, were calculated, and the risk of bias was assessed.

Results: Among the 8,125 unique records identified, 53 articles describing 33 unique systems, were qualitatively, systematically reviewed. Two distinct design categories emerged — simple VH and VH augmented with health sensors and trackers. Of the 53 articles, 16 (26 studies) with 44 primary and 22 secondary outcomes were included in the meta-analysis. Meta-analysis of the 44 primary outcome measures revealed a significant difference between intervention and control conditions, favoring the VH intervention (SMD = .166, 95% CI .039-.292, P=.012), but with evidence of some heterogeneity, I^2=49.3%. There were more cross-sectional (k=15) than longitudinal studies (k=11). The intervention was delivered using a personal computer in most studies (k=18), followed by a tablet (k=4), mobile kiosk (k=2), head-mounted display (k=1), and a desktop computer in a community center (k=1).

Conclusions: We offer evidence for the efficacy of VH in patient-facing systems. Considering that studies included different population and outcome types, more focused analysis is needed in the future. Future studies also need to identify what features of virtual human interventions contribute toward their effectiveness.

XSL•FO
RenderX

Introduction

Patient-facing systems are digital technologies that offer health services and engage people in their health and wellbeing [1]. These systems promote and facilitate both patient engagement with the health care system and patient empowerment in self-care [2]. Globally, one in three adults suffers from multiple chronic conditions [3]; self-management of chronic conditions has become increasingly complex, requiring sophisticated knowledge, motivation, and skill by patients and their families [4]. Engaging in self-management reduces hospitalization and improves health outcomes and quality of life [5]. Patient-facing systems can facilitate self-management of chronic conditions and fill the gap of time and distance to meet with health care providers. Data increasingly support the value of these patient-facing systems, in the form of mobile health (mHealth) applications [6-9] or sensors to monitor physiological parameters (eg, blood glucose) and behaviors (eg, activity) [10-14]. The latest technological addition to patient-facing systems is computer-controlled virtual humans (VH).

Unlike formless conversational agents, like smart speakers or chatbots, VH bring together the capabilities of both a conversational agent and an interactive avatar (computer-represented digital characters). While their humanlike physical appearance is computer-generated (ie, animated), VH are not human-controlled [15-20] nor cartoonlike [21,22]. They are controlled algorithmically, based on active or passive user input during an interaction. These algorithms can simulate key properties of human face-to-face conversation — both verbal and nonverbal (eg, gaze, emotions, head movements, and metaphoric gestures). While humanlike appearance, movements, and nonverbal behavior offer VH the added advantage of communicating empathy and building rapport, they can also bring psychological and social concerns that may not arise when interacting with formless conversational agents [23-26].

Attempts to make computer interfaces anthropomorphic are not new [27,28]. The sophistication of current digital interlocutors, however, can largely be attributed to the recent advancements in artificial intelligence and computer graphics. Recent variants of anthropomorphic interfaces include relational agents, which are agents designed to build and maintain long-term relationships with users [29], including voice-based, intelligent virtual assistants (eg, Amazon's Alexa, Microsoft's Cortana, or Apple's Siri) and text-based chatbots that run on instant messaging applications [30]. Intricate taxonomies capturing different aspects of conversational agents can be found elsewhere [26,31-33].

While intelligent virtual assistants and chatbots have gained mainstream popularity, VH applications are still in their infancy. As graphic rendering capacities progress and ubiquitous computing peripherals such as virtual reality (VR) and augmented reality head-mounted displays (HMDs) become inexpensive, VH will be increasingly adopted for everyday use. Indeed, similar to education and training [34,35], health care researchers and practitioners have already begun to explore the use of VH in health-related assessments and interventions. As VH come of age and stakeholders in health care deliberate whether to adopt this new computing technology, it is crucial that we examine how prior applications have fared in affecting health outcomes.

To our knowledge, VH in patient-facing systems have not been surveyed before. Only recently, other types of conversational agents in health have been reviewed [31,36-40]. Provoost et al [36] reported a scoping review of 54 articles (49 studies) on low-tech embodied conversational agents in clinical psychology. Laranjo et al [37] surveyed conversational agents with unconstrained natural language input capabilities in health care and included 17 articles, of which only 2 were randomized controlled trials (RCTs). Within mental health, both chatbots ([39]; 10 studies) and relational agents ([38]; 13 studies) have been reviewed. Furthermore, the current state of chatbots and embodied conversational agents as expert systems was recently surveyed ([31]; 40 articles). Kocaballi et al [40] reported a systematic review of how conversational agents can be personalized. It included 13 studies and found that personalization features were assessed for user satisfaction, not in improving health outcomes. None of these reviews included a meta-analysis nor focused on VH. The purpose of this study was to conduct a systematic review and meta-analysis of VH in patient-facing systems.

Methods

Overview

This systematic review of the English-language scholarly literature followed standard guidelines for conducting and reporting systematic reviews, including Preferred Reporting Items for Systematic Reviews and Meta-analyses [41] and guidelines from the Cochrane Diagnostic Test Accuracy Working Group [42,43].

Search Strategy

Literature searches were performed from inception to December 31, 2019 in Google Scholar and 7 online databases: MEDLINE, EMBASE, PsycINFO, CINAHL, Cochrane Central Register of Controlled Trials, PubMed, and ACM Digital Library. Search queries covered 3 domains: (1) avatars, (2) (digital) narratives, and (3) virtual assistants (for details, see Multimedia Appendix 1). Our search was limited to peer-reviewed articles published in English. Manual searches were extended to bibliographies of review articles. Figure 1 shows a summary of the literature search.

Figure 1. Summary of the literature search.
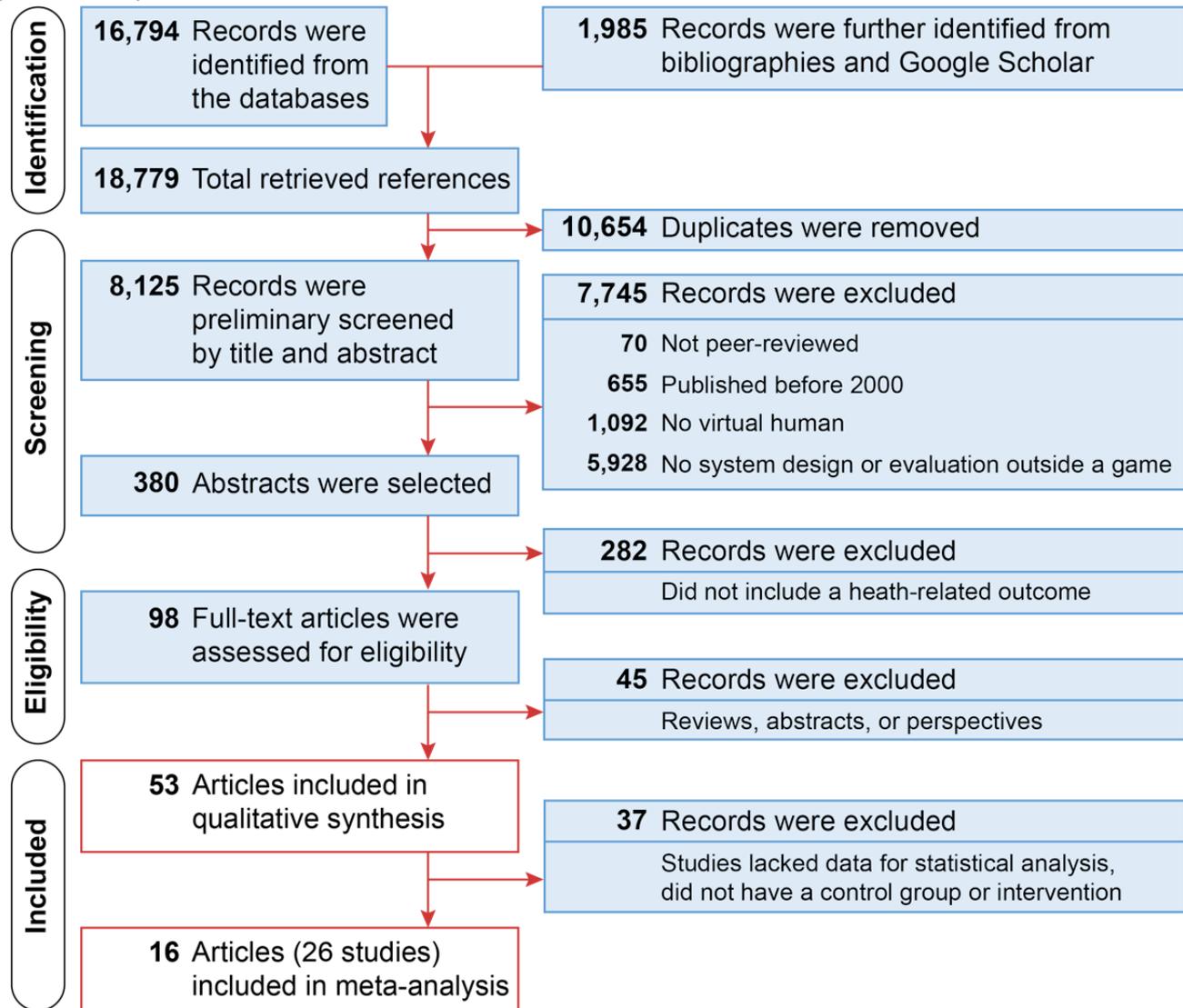
Inclusion and Exclusion Criteria

The topic of our review crosscuts two disciplines — health care and computer science — and their disciplinary priorities are divergent if not orthogonal. On the one hand, health care research prioritizes reporting and replicating empirical evidence of efficacy. On the other hand, driving innovation is a key mission of computer science research. Thus, empirical investigation of these innovative designs — particularly replication of such studies — often takes a back seat. To offer a comprehensive review of the topic at hand, we first present a qualitative, systematic review of VH in patient-facing systems. Some of those articles were then included in a quantitative meta-analysis (Figure 1).

Qualitative Systematic Review

Studies were included in the qualitative review if they met the following criterion: designed or evaluated VH for health-related outcomes in a patient-facing system. Studies were excluded if they used VH for the training and education of health care professionals or students.

Meta-Analysis

Some of the articles included in the qualitative synthesis were further included in a meta-analysis if they met the following criteria: (1) compared the effectiveness of VH in a health-related outcome in a target population against a control group with no VH; (2) studied humans of any age; and (3) reported the sample size and mean and variance of the outcome measure in control and experimental groups. Studies were excluded if they did not use a comparator that was equivalent but different from VH (eg, [44]).

Data Extraction

All records were first downloaded into an EndNote X8.2 library [45], and duplicates were removed. Titles and abstracts were then screened for inclusion and exclusion criteria. To accelerate screening, records were first collated into topics using keyword searches in the EndNote library and then reviewed for inclusion. Two investigators (the first and second authors) independently assessed the relevancy of search results and selected full text articles for further review. The second and third authors independently abstracted the key study factors into a data extraction form and then came to a consensus on which studies

met the inclusion criteria; the first author made the definitive decision when discrepancies would arise.

Data from eligible articles were extracted into a spreadsheet. For the qualitative review, data included target population, design objective, type of evaluation, principal findings, and VH characteristics. For the meta-analysis, studies were summarized using the PICOTS (population, intervention, comparison group, outcomes, time frame, setting) framework [46].

Risk of Bias and Quality Assessment

Articles included in the meta-analysis were assessed for quality and risk of bias using the latest criteria from the Cochrane Consumers and Communication Review Group [47]. Studies were categorized as low risk, unclear, or high risk for each Cochrane bias domain: selection, performance, detection, attrition, reporting, and other [47]. Publication bias was assessed using Egger's test [48] and a P-curve analysis [49] and graphically examined using funnel plots [50].

Statistical Analysis

All statistical analyses were performed in R using the meta [51] and metafor packages [52,53]. Standardized mean differences (SMD) were used as the effect size to quantify the effectiveness of VH. The overall effectiveness was estimated using a random-effects model, and forest plots were used to graphically present the combined effectiveness. Statistical heterogeneity among the studies was assessed by I^2, which estimates the percentage of total variation across studies due to heterogeneity rather than chance alone. We considered heterogeneity statistically significant at $P<.05$ and used random-effects models to take into account the heterogeneity among the included studies. Subgroup analysis was conducted across studies by the type of health outcome — health and wellbeing as well as attitudes toward health and wellbeing.

Results

Qualitative Systematic Review

A total of 16,794 search records were retrieved from the databases, and 1,985 additional records were identified from the bibliographies and Google Scholar. After removing duplicates, we screened the titles and abstracts of 8,125 articles; 380 articles involved a functioning VH system outside a game. Because the computing technology for creating VH did not exist prior to circa 2000, all studies published before 2000 were excluded. These 380 abstracts were further reviewed for their relevance to health, and 282 articles were excluded because the VH was involved in contexts like training, education, or demonstration. The remaining 98 articles underwent full-text review, and 53 articles met the inclusion and exclusion criteria for the qualitative systematic review.

A total of 30 health-related outcomes were identified in the 53 eligible articles, targeting 25 types of populations (Table 1) and 6 modalities of technology delivery apart from a personal computer, desktop, or laptop (Table 2). Unconstrained speech input was sparsely used in VH systems [54,55]. While most systems allowed constrained user input via a touchscreen or keyboard, some systems reacted to nonverbal conversational inputs, such as gaze [56] or proximity [57].

In the 53 eligible articles, 33 unique systems were identified. Of these, 9 systems were used for health assessments and the rest for health interventions (Table 3). Two broad design categories emerged — simple VH and VH with health sensors or trackers (Table 3). These additional trackers did not augment the interaction capabilities of VH, but provided additional data about users (eg, via a heart rhythm monitor, respiration sensor, or eye tracker).

Table 1. In the 53 eligible articles, 30 health-related outcomes and 25 target populations were identified.

Health outcome	Target population	Studies
Improve quality of life	Women with overactive bladder (OAB) symptoms	[58][a]
Self-manage chronic conditions	Individuals with chronic atrial fibrillation (heart condition)	[59-61]
	Individuals with spinal cord injury	[62]
Engage in physical activity	Older adults	[63][a] [64,65]
	Individuals with Parkinson's disease	[66]
	Inactive older adults with low socioeconomic status	[67][a]
	Healthy adults (no reported health conditions)	[68,69][a][70]
	Individuals with schizophrenia	[71]
Improve mood	Individuals with depression	[72][a] [73]
Assess auditory verbal hallucinations (AVH)	Individuals with schizophrenia	[74]
Stress management	Women	[75]
	Individuals with chronic pain and depression	[76][a] [77]
Healthy eating	Women	[75]
	Healthy adults (no reported health conditions)	[69][a]
Improve social skills	Children with autism spectrum disorders (ASD)	[44]
	Individuals with schizophrenia	[78][a]
Assess PTSD[b] symptoms	US military service members	[54][a] [55]
Assess body image disturbance (BID)	Women on diet (nonclinical)	[79]
Anxiety toward death	Older adults	[80]
Find health-related information online	Individuals with low health and computer literacy	[81][a] [82]
Explain health documents	Individuals with low health literacy	[83][a] [65,84-86]
Attitude toward regular physical activity	Healthy adults (no reported health conditions)	[87][a]
Attitude toward breastfeeding	Pregnant women in their third semester	[88][a] [89,90]
Attitude toward weight loss	Healthy adults (no reported health conditions)	[91][a]
Retention of medication knowledge	Individuals with type 2 diabetes mellitus	[92][a]
Attitudes toward prenatal testing for Down syndrome	Nulliparous women	[93]
Improve medication adherence	Individuals with schizophrenia	[71]
Assess emotion recognition	Adults with ASD	[94]
	Individuals with schizophrenia	[95][a]
	Children with ASD	[96]
Preconception risk assessment	Women	[97]
Assess the effects of social rejection	Individuals with psychotic disorder	[98]
Assess social attention	Children with ASD	[99]
Assist in deep breathing	Healthy adults (no reported health conditions)	[100]
Substance use counseling	Individuals with alcohol use disorder	[56]
	Individuals with opioid use disorder	[101]
Patient trust	Healthy adults (no reported health conditions)	[102]
Assess social anxiety disorder	Women with high social anxiety	[103]

Health outcome	Target population	Studies
Alleviate social isolation	Older adults	[57]
Understand the distinction between connective and fatty tissue in the breast	Mammography-eligible middle-aged women (40-74 years old)	[104]
Pill count adherence >80%	HIV-positive African American men who have sex with men	[105]

[a]Studies included in the meta-analysis.

[b]PTSD: post-traumatic stress disorder.

Table 2. Technology characteristics identified in the eligible studies.

Technology characteristics	Studies
Unconstrained speech input	[54,55]
Computer at a community center or school	[44,67]
Smartphone	[59-61]
Head-mounted display (HMD)	[74,78,79,98,99,103]
Virtual reality (VR) in a PC or HMD	[74,78,79,96,98,99,103]
Mobile kiosk with a computer	[63,65,84-86,88]
Tablet	[63,65,66,70,76,77,80,88]

Table 3. Two broad categories of virtual humans emerged from the 53 articles included in the qualitative review.

Type of use	Number of simple virtual humans	Number of virtual humans with health trackers
Intervention	34 [44,57,58,62-64,67-69,71,72,75-78,80-93,97,101,102,104,105]	9 [56,59-61,65,66,70,73,100]
Assessment	7 [54,74,79,94-96,98]	3 [55,99,103]

Virtual Humans for Health Interventions

Of the 34 articles that described simple VH in health-related interventions (Table 3), 27 were based on the same core technology for generating speech [106] and nonverbal behavior [107]. The typical system included a knowledge base of domain-specific top-level dialog fragments (Figure 2). These subdialogs were scripted and then reused to generate natural language speech using a hierarchical transition network, based on augmented transition networks. Augmented transition networks are mathematical structures that can model the grammar of relatively complex natural languages [108]. The generated text was then converted to synthetic speech by the Behavior Expression Animation Toolkit [107]. The Behavior Expression Animation Toolkit also generated appropriate and contextualized nonverbal behavior for the VH. The front end consisted of 3 components: (1) an animated image of the VH communicating with users using speech and gestures, (2) a dynamically updated but constrained multiple-choice menu for user input (via touch or keyboard), and (3) other content supporting the conversation as necessary, like text and images. VH in such a system talked to users with synthetic speech and presented synchronized nonverbal behavior, such as hand gestures and eyebrow raises for emphasis, looking away to signal turn-taking, and posture shifts to mark topic boundaries.

The VH interface in Figure 2 was mostly delivered via smartphones or personal computers. By tailoring the knowledge base, VH were designed for a range of domains, from managing depression [77], through end-of-life planning [80], stress management [75], and educating about breastfeeding [89], to preconception care [97].

Design approaches were varied in the rest of the studies. Dworkin et al [105] followed the same concepts as in Figure 2 but did not discuss implementation details. Two studies used a third-party system [44,72]. Another used VR and an HMD to deliver the VH intervention [78]. Some studies did not detail exactly how the conversations between humans and VH were automated [58,68,78,92], other than mentioning that nonverbal behaviors were programmed to align with verbal utterances. Two studies modeled their intervention after tutorials, with functionality, such as play, pause, repeat, and answering multiple-choice questions [58,92].

Some designs augmented VH with sensor-based tracking (Table 3), such as heart rhythm monitors for managing chronic heart failure [59-61], a pedometer for promoting walking [65,66], and a breathing monitor for assisting in meditation [100].

Figure 2. The most common structure of a simple virtual human system designed for health-related interventions. BEAT: Behavior Expression Animation Toolkit.
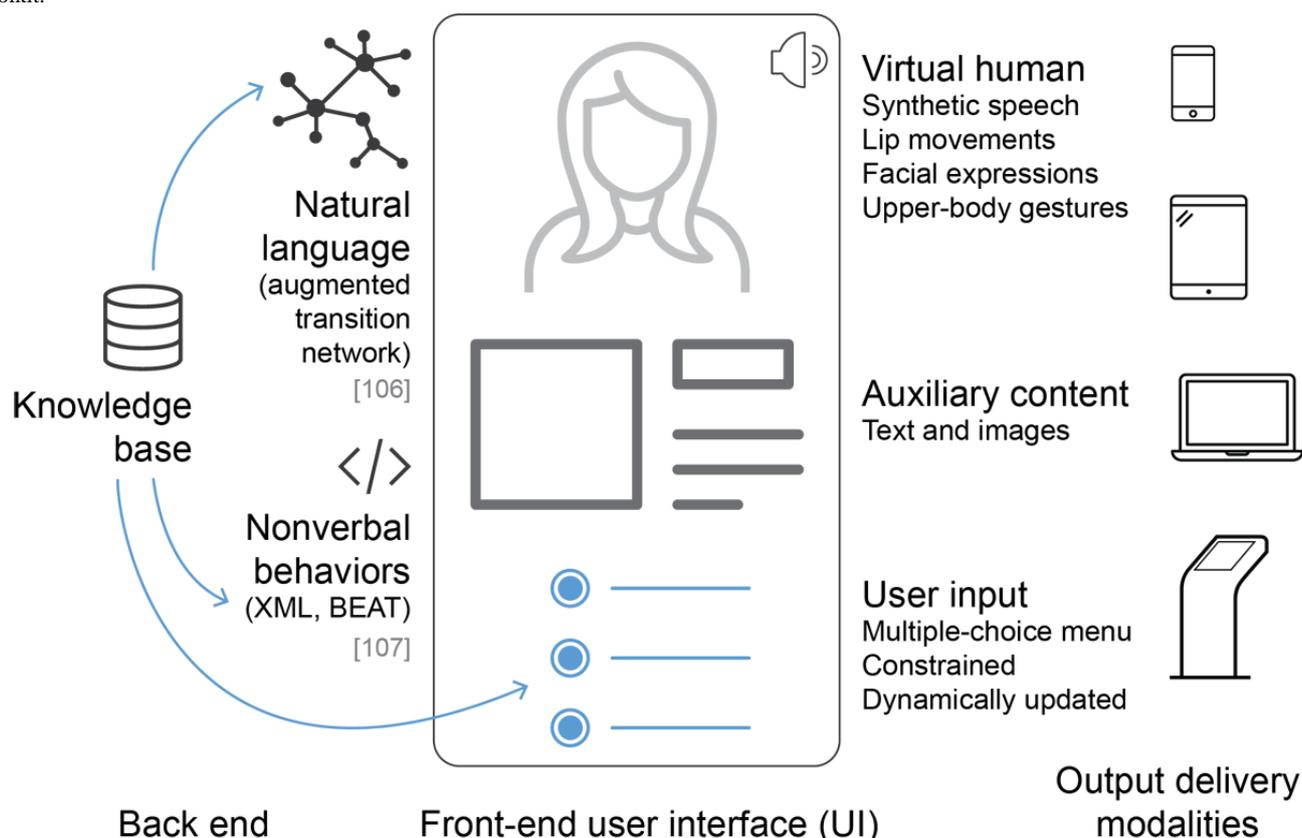
Virtual Humans for Health Assessments

Compared with interventions, fewer studies were found for health-related assessments (Tables 1 and 3); most of these were for people with autism spectrum disorder [94,96,99] and psychotic disorders [74,98,103]. Other studies used VH to assess PTSD symptoms [54,55], emotion recognition [95], and body image disorder [79]. Six studies used VH in a VR environment; 5 used HMDs [74,79,98,99,103], while 1 used a computer [96]. Two studies used advanced tracking systems; these tracking modules were more sophisticated than those used in health interventions. To assess social attention in high-functioning children with autism spectrum disorder, researchers used a sensor to measure head orientation and rotational motion [99]. The MultiSense framework was used for the SimSensei Kiosk [55], which tracks an array of perceptual signals, like smile intensity, gaze direction, and lack of facial expressions, and allows unconstrained speech input [54].

Virtual Human Characteristics

Overall, the physical appearances of VH were primarily created using 3-dimensional (3D) character modeling and animation software, such as the Unity3D game engine. They were designed to be racially ambiguous [64] or accordant with target users [86]. But Marcos-Pablos et al [95] generated dynamic expressions for VH (eg, anger, happiness, sadness) by first using a laser scanner to obtain a 3D face model and then using animation software. The speech of VH was either synthetically produced [76] or lip-synced to audio-recorded voice narrations [58,92]. Synthetic speech was generated using commercially available software.

Characteristics, personalities, and mannerisms of VH were manipulated to build rapport with end users [102]. Schulman et al [87] explored the use of social dialog (small talk) and persuasive arguments (about 30 turns) to change attitude toward exercise. VH argued about the advantages of regular exercise and against statements that emphasized disadvantages. Friederichs et al [68] used social dialog at the beginning of an intervention and during transitions between different parts. To persuade users, Andrade et al [58] developed a peer avatar — based on the front-face view of the participant — who delivered a tutorial about self-managing overactive bladder symptoms. Motivational interviewing techniques were used by Friederichs et al [68] to promote physical activity in adults.

Another study explored the use of personal stories available on the internet to personalize the VH's message and change health behavior [91]. These stories were indexed based on participants's personal state of change during weight loss. To promote breastfeeding, lactation education was framed from a feminist perspective [90]; a VH was designed to motivate women in their third semester to follow breastfeeding recommendations by the US Centers for Disease Control and Prevention. This was realized by introducing a feminist introductory and closing script (eg, "I like to think of myself as a fairly progressive woman, and I hope I do not offend you by some of my opinions" or "A lot of people think that breastfeeding represents a dilemma for feminism.") [90]. Participants who did not self-identify as feminists were significantly less satisfied with the feminist agent than the controls.

Middle-aged Caucasian and African American VH were designed to achieve racial concordance with users [92], but racial concordance did not significantly affect perceived similarity [86]. Researchers also investigated the effects of attire, background image, and alignment on trust. Study participants found a patient-aligned VH explaining an informed consent document more trustworthy than a medical-aligned or federal-aligned VH. Deictic gestures, such as pointing to a document with a finger or open hand, were designed to aid in document explanation [85]. When explaining a medical document, study participants with low literacy were more satisfied with such a VH than a human. Finally, a common theme across many VH designed for health-related interventions was continuity of care, or longitudinal engagement with end users.

Design Process

Of the 53 articles, 23 explicitly described their VH design process [55,58,59,68,71,73,75,78,80,82-87,89-93,95,97,104]. The most common design approaches were collaboration with domain experts, qualitative observation of similar human-human interactions, and adoption of public health and governmental guidelines. Some studies conducted focus groups and one-to-one interviews with target users [59,68,90,92,97,104]. Others informed their design decisions from prior literature or theoretical frameworks [58,59,68,71,90,91,93]. Qualitative observations were often videotaped and coded to generate conversation content [55,71,84]. VH for health-related assessments were always designed in collaboration with health care providers but not end users (Table 3).

Theory

Only a few papers explicitly mentioned adopting theoretical frameworks to ground their design of VH [56,60,68,70,91,93,100]. The chronic care model was used to design VH for people with chronic illnesses [109]; the following 3 guidelines were adopted: facilitate communication between patients and providers, make patients aware of the latest care guidelines consistent with scientific evidence, and motivate patients to manage their health [60].

Two frameworks of behavior change were used widely — the transtheoretical model (TTM) of health behavior change [110] and motivational interviewing (MI) counseling style [111]. The transtheoretical model operationalizes intentional behavior change with 3 core constructs — stages of change (precontemplation, contemplation, preparation, action, and maintenance), processes of change, and levels of change — and recommends stage-matched interventions. VH could ask questions to determine these stages and offer matched interventions, such as therapeutic alliance [56,70,91]. MI was developed to elicit people's readiness to change, explore their attitudes toward change, and transition ambivalence toward reducing resistance to behavior change. The conversation style of VH was sometimes designed based on the MI technique [56,68]. One study designed VH to mirror basic user actions [112], like inhaling and exhaling, to promote behavior change [100]. In another study [93], VH played the role of a decision coach to facilitate shared decision making [113,114]. They informed users of all available options, provided detailed information about those options, and finally assisted in choosing one of them.

Design Guidelines

Four papers explicitly offered design guidelines for VH [57,60,64,71]. To design VH that can maintain long-term social relationships with end users (ie, relational agents), Bickmore et al [64] recommended developing a model of user-agent relationship; using relational behavior like social dialog, empathy, and humor; maintaining a memory of past interactions; and providing some variability in agent behavior and overall variability in the dialog structure. When designing VH for mental health interventions that are also relational agents, special design considerations may be needed. For patients with schizophrenia, Bickmore et al [71] recommended prolonging the introduction and conclusion phases of the conversation, using clear and concrete language that focuses on real events to reinforce reality, and not relying solely on nonverbal behaviors like a gaze-away gesture. Specific populations, health outcomes, and device modalities may need additional design considerations. When addressing loneliness in older adults, Ring et al [57] recommended VH engage in social interactions, such as small talk and games, help older adults stay connected with friends and family via contemporary technologies, and alleviate mood disorders or loneliness through different types of talk therapy. When designing for smartphones, Bickmore et al [60] recommended keeping interactions short, allowing interruptions during interactions, and using constrained user input.

Meta-Analysis

A total of 26 studies (16 articles) published between 2000 and December 31, 2019 were eligible for the meta-analysis, targeting 11 types of populations and including 10 studies with healthy adults [68,69,81,83,87], 3 studies with older adults [63,67], 3 studies with women [58,88], 2 studies with individuals with schizophrenia [78,95], and 3 studies with individuals with depression [72,76].

Study Characteristics

The included studies comprised approximately 1400 participants across 13 health and wellbeing objectives. The PICOTS information [46] from the 26 studies is available in Multimedia Appendix 2. Outcomes included 44 primary outcomes and 22 secondary outcomes. Of the 26 studies, 9 focused on attitudes toward health and wellbeing [81,83,87,88,91]. There were more cross-sectional (*k*=15) [54,68,81,83,87,88,91,92,95] than longitudinal studies (*k*=11) [58,63,67,69,72,76,78,88]. Longitudinal studies ranged from 1 month to 6 months (see Multimedia Appendix 2). The VH intervention was delivered using a personal computer in most studies (*k*=18), followed by a tablet (*k*=4), mobile kiosk (*k*=2), HMD (*k*=1), and a desktop computer in a community center (*k*=1).

Evidence of Intervention Efficacy

As evident from Table 2 and Multimedia Appendix 2, the 26 studies eligible for meta-analysis varied in terms of outcomes, target population, timing, and intervention design. Hence, a random-effects model was used. A random-effects model does not assume that the estimated effects come from a single homogeneous population, but that true effect sizes vary from

study to study. Hedges' *g* was calculated for each reported outcome, and SMD were used as the effect size to quantify the overall evidence for VH interventions. Next, we report the meta-analysis of the 44 primary outcomes across 26 studies. A meta-analysis of all 66 outcomes (including 22 secondary outcomes) is available in Multimedia Appendix 2.

The meta-analysis of data from 26 studies (44 outcomes) revealed a significant difference between intervention and control conditions, favoring the VH intervention (SMD .166, 95% CI .039-.292, 95% prediction interval –.548 to .879, *P*=.012) but with evidence of some heterogeneity: $I^2 = 49.3\%$, 95% CI 28.1%-64.3% (Figure 3).

A 3-level model (level 2: different outcome measures; level 3: different studies) did not capture a significant amount of variability in the data (*P*>.05). Thus, a 2-level model was used. The between-study heterogeneity of the data was moderate: τ^2=.12, I^2=49.3%. We examined whether this heterogeneity was caused by outliers or influential cases [115,116]. No influential cases were detected, but we spotted 3 outliers [67,76,95]. While

2 of these outlier studies found extremely positive evidence about VH efficacy [67,76], 1 study found extremely negative evidence for one of its outcome measures: happiness emotion recognition by patients with schizophrenia [95]. However, the meta-analysis after removing these outliers (*k*=41, SMD .144, 95% CI .028-.260, *P*=.016; I^2=33.7%, 95% CI 2.6%-54.8%) was similar to the original analysis (Figure 3). Because these 3 studies were interesting outliers, but neither error nor influential outliers [117], we chose to retain them.

A subgroup analysis for health-related outcomes and health-related attitudes was conducted, but no significant difference was found in the overall effect between outcome types (*P*=.762). The number of studies was insufficient to conduct subgroup analyses for different population or outcome types.

To explore publication bias, a funnel plot was generated. Egger's test was not significant (*P*=.70), indicating no substantial asymmetry in the funnel plot. Furthermore, the *P*-curve analysis did not indicate publication bias (Figure 4).

Figure 3. Forest plot of the meta-analysis of health-related virtual human interventions from 26 studies (44 primary outcomes). a-PDHA: anonymized post-deployment health assessment; ACT: physical activity; BDI-2: Beck Depression Inventory-II; BICEP: brief informed consent evaluation protocol; DAS–SF2: Dysfunctional Attitude Scale-Short Form 2; DIET: fruit and vegetable consumption; EQ–5D–5L VAS: 5-level version of the EuroQol 5D visual analogue scale; FVS: NIH/NCI Fruit and Vegetable Scan; HRQOL: health-related quality of life; OABq: overactive bladder questionnaire; PDHA: post-deployment health assessment; PTSD: post-traumatic stress disorder; QIDS–SR: Quality of Life Enjoyment and Satisfaction Questionnaire-Short Form; SBS: social behavior scales; SMD: standardized mean difference; SVH: social virtual human.

Study	Outcome Measures	SMD	95%–CI
[57]	OABq HRQOL	0.67	[0.04; 1.30]
[82]	BICEP; vs. human	0.15	[−0.77; 1.08]
[82]	BICEP; vs. self	0.98	[0.04; 1.91]
[68]	ACT step count	0.00	[−0.52; 0.52]
[68]	DIET step count	−0.01	[−0.54; 0.53]
[68]	ACT + DIET step count	−0.00	[−0.54; 0.53]
[68]	ACT FVS score	0.21	[−0.31; 0.73]
[68]	DIET FVS score	0.54	[−0.01; 1.08]
[68]	ACT + DIET FVS score	0.34	[−0.20; 0.88]
[62]	Step count	0.12	[−0.23; 0.47]
[80]	Find a correct clinical trial	0.24	[−0.18; 0.67]
[71]	BDI−2	0.47	[−0.41; 1.35]
[71]	QIDS−SR	0.07	[−0.79; 0.94]
[71]	DAS−SF2	1.05	[0.13; 1.98]
[71]	EQ−5D−5L VAS	−0.62	[−1.50; 0.27]
[87]	Attitude toward breastfeeding at third triemster	0.14	[−0.87; 1.16]
[87]	Attitude toward breastfeeding at hospital discharge	−0.33	[−1.43; 0.77]
[67]	Days with physical activity	−0.06	[−0.28; 0.16]
[66]	Change in walking	1.21	[0.52; 1.89]
[54]	PTSD symptoms reported; vs. a−PDHA (including veterans)	0.14	[−0.11; 0.38]
[54]	PTSD symptoms reported; vs. PDHA	0.55	[−0.03; 1.12]
[54]	PTSD symptoms reported; vs. a−PDHA	0.46	[−0.11; 1.04]
[90]	Weight loss self efficacy for indexed stories	0.47	[−0.04; 0.99]
[90]	Weight loss self efficacy for random stories	−0.03	[−0.64; 0.58]
[90]	Weight loss decisional balance for indexed stories	0.02	[−0.49; 0.52]
[90]	Weight loss decisional balance for random stories	0.35	[−0.26; 0.96]
[94]	Emotion recognition: anger	−0.51	[−1.14; 0.12]
[94]	Emotion recognition: happiness	−0.68	[−1.32; −0.05]
[94]	Emotion recognition: sadness	0.74	[0.10; 1.39]
[94]	Emotion recognition: fear	−0.11	[−0.73; 0.52]
[94]	Emotion recognition: surprise	0.54	[−0.09; 1.17]
[77]	SBS	−0.25	[−0.74; 0.25]
[77]	Vocal skill	−0.02	[−0.51; 0.47]
[77]	Nonverbal skill	−0.27	[−0.76; 0.22]
[77]	Conversational skill	−0.24	[−0.73; 0.25]
[91]	Medication knowledge recall; vs. text	−0.09	[−0.59; 0.42]
[91]	Medication knowledge recall; vs. voice only	−0.15	[−0.65; 0.36]
[91]	Medication knowledge recall; vs. static image	−0.15	[−0.66; 0.35]
[86]	Attitudes toward exercise, VH vs text	0.55	[−0.31; 1.42]
[86]	Attitudes toward exercise, VH vs. text + social	0.82	[−0.08; 1.72]
[86]	Attitudes toward exercise, SVH vs text	−0.51	[−1.30; 0.27]
[86]	Attitudes toward exercise, SVH vs text + social	−0.27	[−1.05; 0.52]
[75]	Stress management behaviors at 9 weeks	0.63	[0.31; 0.96]
[75]	Stress management behaviors at 12 weeks	0.57	[0.25; 0.90]

⟵— no virtual human better   virtual human better —⟶

Random effects model   0.17 [0.04; 0.29]
Prediction interval   [−0.55; 0.88]
Heterogeneity: I^2 = 49%

−1   0   1

Figure 4. The observed *P*-curve has an estimated power of 99% (left and right), significant right skewness, $P_{full} < .0001$, $P_{half} < .0001$ (left), and no significant flatness, $P_{full} > .9999$, $P_{half} > .9999$.
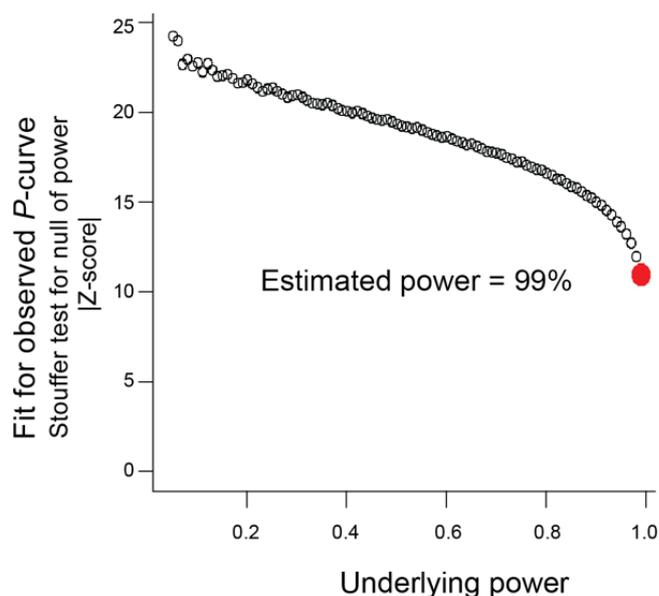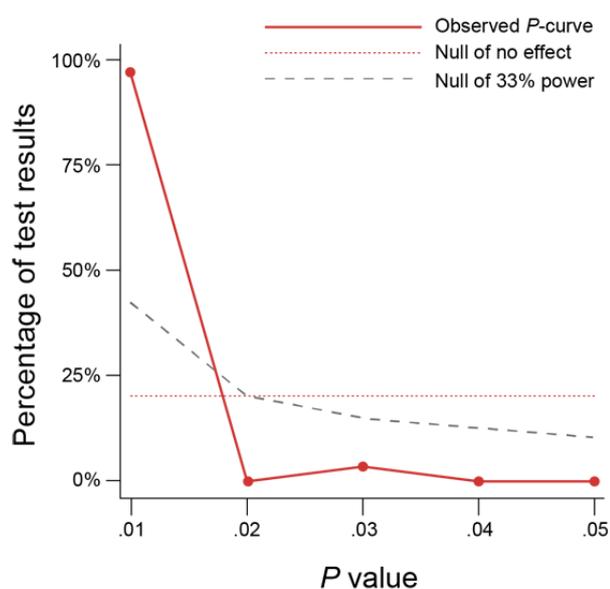
Risk of Bias in Included Studies

The quality of studies included in the meta-analysis was evaluated for risk of bias (Figures 5 and 6). Randomization was adequate in 19 studies [58,63,67-69,72,78,87,88,91,92,95] but unclear in 7 others [54,76,81,83]. Allocation was reported to be concealed in 11 studies [58,63,67,72,88,92,95], but other studies did not provide enough information to assess allocation bias. It was unclear in 15 studies whether participants and research personnel were blinded to the allocated interventions [54,68,69,76,78,81,83,87,91].

Eleven studies were successful in blinding participants and research personnel to the allocated interventions

[58,63,72,81,87,92,95], while 5 studies were not [54,78,88]; the remaining studies were unclear. Blinding the assessors of outcomes was achieved in 8 studies [58,63,78,92,95], 2 studies did not blind research personnel [67,72], and 16 did not clearly report this [54,68,69,76,81,83,87,88,91]. Attrition was high (>20%) in 5 studies [63,68,72,78]. With respect to selective reporting, we identified 13 studies that did not report either descriptive statistics for nonsignificant outcomes or participant demographics [54,67-69,76,88,92]. In assessing other potential sources of bias, we identified 11 studies at a risk of bias due to a small sample size [87,88] or self-selection from provider invitations [72] and websites like Craigslist or Amazon Mechanical Turk [54,69,87,91].

Figure 5. Summary of the authors' consensus judgment about the risk of bias for each study included in the meta-analysis, by various sources of potential bias.
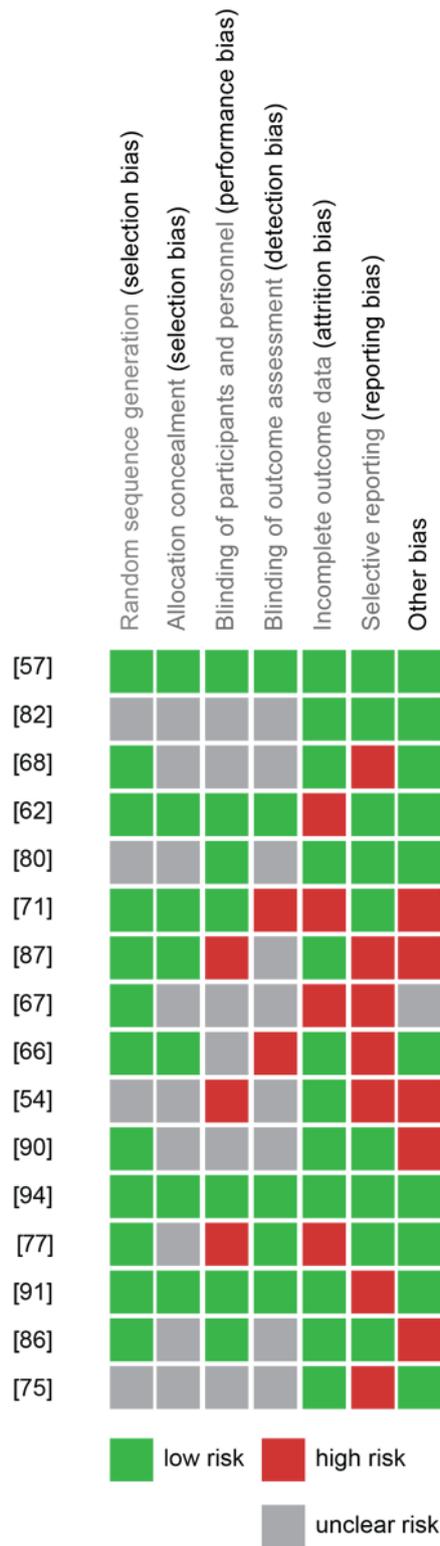
Figure 6. Risk of bias presented as percentages across all 26 studies included in the meta-analysis.
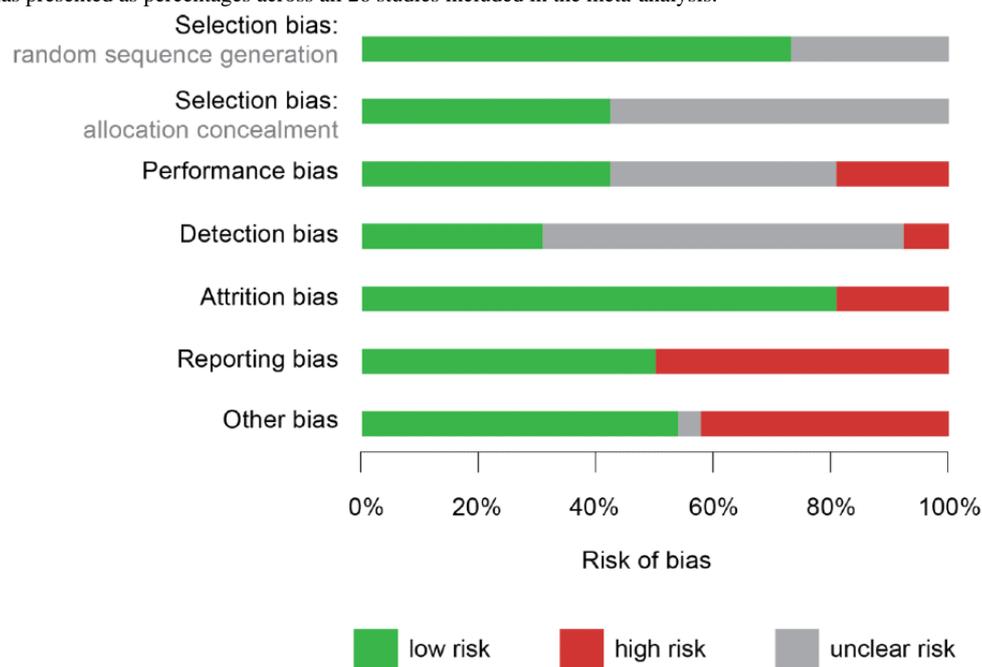
Discussion

Primary Findings

This meta-analysis found VH interventions significantly more effective than other types of traditional interventions that did not include conversational agents (SMD .166, 95% CI=.039-.292, P=.012). This is an important finding because effective interaction modalities are needed to promote improved consumer or patient engagement in health, thereby promoting behavior change and the management of chronic conditions. The effect size may have been small because the studies involved many different population and outcome types. As more studies become available in the future, some populations and outcome types may show stronger effect sizes than others.

The efficacy of VH may also depend on the type of intervention for which they are used, such as delivering cognitive behavioral therapy vs delivering education. However, the number of current studies was insufficient to conduct such subgroup analyses. Nevertheless, the effectiveness of VH did not significantly differ between health-related outcomes and health-related attitudes.

Like prior reports [37], we found considerable heterogeneity in evaluation methods and outcome measures and a predominance of quasiexperimental study designs or usability studies over RCTs. But unlike previous reviews [36], we found most applications were fully developed, and many of them were either evaluated or soon to be evaluated in RCTs. Our findings cannot be compared directly with the previous reviews because each of the reviews used a different definition of conversational agents or avatars. Typically, earlier reviews did not discriminate between conversational agents with and without a humanlike physical appearance; therefore, they did not discriminate between whether they were able or unable to engage in nonverbal conversation. Only 5 of the 26 studies that were included in the meta-analysis compared the efficacy of VH interventions with text-based chatbots or any other type of

conversational agents (Multimedia Appendix 2) [58,68,87,91,92]. Other studies compared interventions against treatment as usual, such as a therapist visit or standard educational materials. Voice-based and text-based conversational agents, without a physical humanlike appearance, were used as comparators in 2 studies [58,92] and 4 studies [68,87,91,92], respectively. Thus, it is too early to tell how much the efficacy of VH in health-related interventions can be attributed to their humanlike physical appearance.

Of course, some health applications, especially some mental health assessments, would only work with VH and could not be replaced with a voice-based or text-based conversational agent. For example, VH were used to assess the influence of unusual voices on daily activities of hallucinating patients [74] and the effect of social rejection in individuals with psychotic disorder [98]. But apart from these, when should a VH intervention be more desirable than other types of conversational agents? Engineering health-related VH interventions is far from trivial and may further cause unintended effects, such as a lack of empathy or a sense of complacency among users [23,24]. When is it worth investing in designing VH over chatbots or smart speakers? This question kept resurfacing in our analysis but remained unanswered. As more studies become available, subgroup analyses of specific interventions could answer this question.

Currently, the prevalence of VH in health applications appears to lag behind that in other areas, such as education and training [118,119] or serious games [120], which is not surprising given the concerns about patient safety [121], ethical and legal issues [122], and perception of professional counselors [123]. Technological advancements will continue to augment VH capabilities; equally important is identifying the design tradeoffs associated with those capabilities in patient-facing systems. When should a conversational agent have a humanlike appearance or any physical embodiment? Would allowing for unconstrained speech and gesture input improve intervention

effectiveness or reduce patient safety and privacy? When are nonverbal cues significant in designing a conversational agent? Does age, gender, socioeconomic status, literacy levels, or certain characteristics of target users mediate the effectiveness of VH interventions?

Finally, we found that the input and output of VH systems have evolved significantly over the last 2 decades, drawing on the most recent technological advancements. While systems in the 2000s extensively used desktops, kiosks, 2-dimensional graphics, and constrained text input [64], more recent systems were developed in 3D, were delivered in VR, sometimes used HMDs, and allowed for unconstrained input, such as speech, gestures, and facial emotions. Two key design trends were identified: (1) multimodal sensing of the user's state using computer vision algorithms and ubiquitous computing technologies, ranging from upper-body gestures to heart rate (input), and (2) striving toward high fidelity, humanlike appearance and behavior of the VH to improve presence and immersion (output).

Limitations and Future Research

The limitations of this review should be noted and can be addressed by future studies. First, not all studies on VH in patient-facing systems were included in our work. This is because they did not present sufficient quantitative information, only reported usability metrics, or did not clarify whether their avatar technology was computer-controlled or human-controlled. Including additional studies and VH designs could allow reinforcing the results reported here or provide different results. Second, some of the studies included in our meta-analysis had relatively small sample sizes (<20 participants); thus, additional caution is recommended when generalizing these results. Third, there was moderate heterogeneity among trials in the meta-analysis (τ^2=.12, I^2=49.3%.). This can be attributed to the different health outcome measures, population types, and health objectives. For example, health objectives ranged from increasing physical activity and improving mood to improving social skills. When more studies become available with the same or similar health objectives, it would be worth updating this study with new results.

Although research on conversational agents began circa 2000, their design and capabilities have changed and diverged substantially as new technologies and sensors have emerged. This change is expected to continue. Future studies are suggested to consider the difference between different types of conversational agents when synthesizing or generalizing the agents. For example, does a physical appearance or nonverbal behavior increase the effectiveness of a conversational agent? In what kind of tasks? Furthermore, there is rich literature on behavior change and health behavior change theories. However, theoretical frameworks explicating how different features of VH work together in building (or disrupting) rapport with patients is lacking. As such models emerge, future studies will need to examine those relationships between model constructs with methods such as meta-analytic structural equation models.

Conclusion

VH are conversational agents with a humanlike physical appearance; autonomy in verbal and nonverbal behavior; and speech, gaze, or gesture interaction capabilities. In patient-facing systems, they can demonstrate listening and empathy, as well as tailor to various sociocultural backgrounds, languages, and literacy levels. We surveyed the existing literature on VH in patient-facing systems — from inception to December 2019. Of the 53 articles reviewed, a meta-analysis of 26 studies with more than 1400 participants showed that VH interventions significantly improve health outcomes compared with other traditional intervention methods. But whether a physical embodiment is crucial for a conversational agent to significantly improve health-related outcomes remains to be explored, as does any effect of the VH's physical appearance, type of voice, or quality of movements.

Although not yet comparable to computer-animated films or high-end video games, the appearance and behavior of VH in health care are increasingly becoming sophisticated, with studies finding that users prefer more humanlike VHs in health care [124,125]. Elsewhere, studies continue to report the possibility of unintended negative user reactions (ie, the uncanny valley effect) when interacting with VH, owing to a mismatch in the levels of realism, either physical, behavioral, or both [23,24,126]. However, the literature on VH in patient-facing systems has not yet examined whether the uncanny valley effect affects patient perception and, in turn, health outcomes.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Literature search details.
[PDF File (Adobe PDF File), 94 KB-Multimedia Appendix 1]

Multimedia Appendix 2

PICOTS information.

[PDF File (Adobe PDF File), 1419 KB-Multimedia Appendix 2]

References

1. Jensen RE, Gummerson SP, Chung AE. Overview of Patient-Facing Systems in Patient-Reported Outcomes Collection: Focus and Design in Cancer Care. JOP 2016 Oct;12(10):873-875. [doi: 10.1200/jop.2016.015685]

2. Ahern DK, Woods SS, Lightowler MC, Finley SW, Houston TK. Promise of and potential for patient-facing technologies to enable meaningful use. Am J Prev Med 2011 May;40(5 Suppl 2):S162-S172. [doi: 10.1016/j.amepre.2011.01.005] [Medline: 21521591]

3. Hajat C, Stein E. The global burden of multiple chronic conditions: A narrative review. Prev Med Rep 2018 Dec;12:284-293 [FREE Full text] [doi: 10.1016/j.pmedr.2018.10.008] [Medline: 30406006]

4. Schulman-Green D, Jaser SS, Park C, Whittemore R. A metasynthesis of factors affecting self-management of chronic illness. J Adv Nurs 2016 Jul 19;72(7):1469-1489 [FREE Full text] [doi: 10.1111/jan.12902] [Medline: 26781649]

5. Riegel B, Lee CS, Albert N, Lennie T, Chung M, Song EK, et al. From Novice to Expert. Nursing Research 2011;60(2):132-138. [doi: 10.1097/nnr.0b013e31820978ec]

6. Singh K, Drouin K, Newmark LP, Filkins M, Silvers E, Bain PA, et al. Patient-Facing Mobile Apps to Treat High-Need, High-Cost Populations: A Scoping Review. JMIR Mhealth Uhealth 2016 Dec 19;4(4):e136 [FREE Full text] [doi: 10.2196/mhealth.6445] [Medline: 27993761]

7. Wakefield BJ, Turvey CL, Nazi KM, Holman JE, Hogan TP, Shimada SL, et al. Psychometric Properties of Patient-Facing eHealth Evaluation Measures: Systematic Review and Analysis. J Med Internet Res 2017 Oct 11;19(10):e346 [FREE Full text] [doi: 10.2196/jmir.7638] [Medline: 29021128]

8. Mendiola MF, Kalnicki M, Lindenauer S. Valuable features in mobile health apps for patients and consumers: content analysis of apps and user ratings. JMIR Mhealth Uhealth 2015 May 13;3(2):e40 [FREE Full text] [doi: 10.2196/mhealth.4283] [Medline: 25972309]

9. Vaghefi I, Tulu B. The Continued Use of Mobile Health Apps: Insights From a Longitudinal Study. JMIR Mhealth Uhealth 2019 Aug 29;7(8):e12983 [FREE Full text] [doi: 10.2196/12983] [Medline: 31469081]

10. Baron KG, Duffecy J, Richardson D, Avery E, Rothschild S, Lane J. Technology Assisted Behavior Intervention to Extend Sleep Among Adults With Short Sleep Duration and Prehypertension/Stage 1 Hypertension: A Randomized Pilot Feasibility Study. J Clin Sleep Med 2019 Nov 15;15(11):1587-1597. [doi: 10.5664/jcsm.8018] [Medline: 31739848]

11. Garg SK, Shah VN, Akturk HK, Beatson C, Snell-Bergeon JK. Role of Mobile Technology to Improve Diabetes Care in Adults with Type 1 Diabetes: The Remote-T1D Study iBGStar in Type 1 Diabetes Management. Diabetes Ther 2017 Aug 29;8(4):811-819 [FREE Full text] [doi: 10.1007/s13300-017-0272-5] [Medline: 28555339]

12. Melin M, Hägglund E, Ullman B, Persson H, Hagerman I. Effects of a Tablet Computer on Self-care, Quality of Life, and Knowledge. The Journal of Cardiovascular Nursing 2018;33(4):336-343. [doi: 10.1097/jcn.0000000000000462]

13. Rathbone AL, Prescott J. The Use of Mobile Apps and SMS Messaging as Physical and Mental Health Interventions: Systematic Review. J Med Internet Res 2017 Aug 24;19(8):e295 [FREE Full text] [doi: 10.2196/jmir.7740] [Medline: 28838887]

14. Wang K, Varma DS, Prosperi M. A systematic review of the effectiveness of mobile apps for monitoring and management of mental health symptoms or disorders. J Psychiatr Res 2018 Dec;107:73-78. [doi: 10.1016/j.jpsychires.2018.10.006] [Medline: 30347316]

15. Fox J, Bailenson J, Binney J. Virtual Experiences, Physical Behaviors: The Effect of Presence on Imitation of an Eating Avatar. Presence: Teleoperators and Virtual Environments 2009 Aug;18(4):294-303. [doi: 10.1162/pres.18.4.294]

16. Magnenat-Thalmann N, Kasap Z. Virtual Humans in Serious Games. 2009 Presented at: International Conference on CyberWorlds; 2009; Bradford, UK. [doi: 10.1109/cw.2009.17]

17. du Sert OP, Potvin S, Lipp O, Dellazizzo L, Laurelli M, Breton R, et al. Virtual reality therapy for refractory auditory verbal hallucinations in schizophrenia: A pilot clinical trial. Schizophr Res 2018 Jul;197:176-181. [doi: 10.1016/j.schres.2018.02.031] [Medline: 29486956]

18. Craig TK, Rus-Calafell M, Ward T, Leff JP, Huckvale M, Howarth E, et al. AVATAR therapy for auditory verbal hallucinations in people with psychosis: a single-blind, randomised controlled trial. The Lancet Psychiatry 2018 Jan;5(1):31-40. [doi: 10.1016/s2215-0366(17)30427-3]

19. Leff J, Williams G, Huckvale MA, Arbuthnot M, Leff AP. Computer-assisted therapy for medication-resistant auditory hallucinations: proof-of-concept study. Br J Psychiatry 2013 Jun 02;202(6):428-433. [doi: 10.1192/bjp.bp.112.124883] [Medline: 23429202]

20. Fabio RA, Caprì T, Iannizzotto G, Nucita A, Mohammadhasani N. Interactive Avatar Boosts the Performances of Children with Attention Deficit Hyperactivity Disorder in Dynamic Measures of Intelligence. Cyberpsychol Behav Soc Netw 2019 Sep 01;22(9):588-596. [doi: 10.1089/cyber.2018.0711] [Medline: 31441667]

21. Ogle AD, Graham DJ, Lucas-Thompson RG, Roberto CA. Influence of Cartoon Media Characters on Children's Attention to and Preference for Food and Beverage Products. J Acad Nutr Diet 2017 Feb;117(2):265-270.e2 [FREE Full text] [doi: 10.1016/j.jand.2016.08.012] [Medline: 27793520]

22. Sinor MZ. Comparison between conventional health promotion and use of cartoon animation in delivering oral health education. Intl J Humanities and Social Science 2011;1(3):169-174.

23. Chattopadhyay D, MacDorman KF. Familiar faces rendered strange: Why inconsistent realism drives characters into the uncanny valley. J Vis 2016 Sep 01;16(11):7 [FREE Full text] [doi: 10.1167/16.11.7] [Medline: 27611063]

24. Dai Z, MacDorman K. The doctor's digital double: how warmth, competence, and animation promote adherence intention. PeerJ Computer Science 2018;4:e168. [doi: 10.7717/peerj-cs.168]

25. Zhou MX, Mark G, Li J, Yang H. Trusting Virtual Agents. ACM Trans. Interact. Intell. Syst 2019 Apr 25;9(2-3):1-36. [doi: 10.1145/3232077]

26. Feine J, Gnewuch U, Morana S, Maedche A. A Taxonomy of Social Cues for Conversational Agents. International Journal of Human-Computer Studies 2019 Dec;132:138-161. [doi: 10.1016/j.ijhcs.2019.07.009]

27. Weizenbaum J. ELIZA---a computer program for the study of natural language communication between man and machine. Commun. ACM 1966;9(1):36-45. [doi: 10.1145/365153.365168]

28. Cassell J. Embodied conversational agents: representation and intelligence in user interfaces. AI magazine 2001;22(4):67-67. [doi: 10.7551/mitpress/2697.001.0001]

29. Bickmore TW, Picard RW. Establishing and maintaining long-term human-computer relationships. ACM Trans. Comput.-Hum. Interact 2005 Jun;12(2):293-327. [doi: 10.1145/1067860.1067867]

30. Hill J, Randolph Ford W, Farreras IG. Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations. Computers in Human Behavior 2015 Aug;49:245-250. [doi: 10.1016/j.chb.2015.02.026]

31. Montenegro JLZ, da Costa CA, da Rosa Righi R. Survey of conversational agents in health. Expert Systems with Applications 2019 Sep;129:56-67. [doi: 10.1016/j.eswa.2019.03.054]

32. Gnewuch U, Morana S, Heckmann C, Maedche A. Designing Conversational Agents for Energy Feedback. 2018 Presented at: International Conference on Design Science Research in Information Systems and Technology; 2018; Chennai, India p. 18-33. [doi: 10.1007/978-3-319-91800-6_2]

33. Diederich S, Brendel AB, Kolbe LM. Designing Anthropomorphic Enterprise Conversational Agents. Bus Inf Syst Eng 2020 Mar 10;62(3):193-209. [doi: 10.1007/s12599-020-00639-y]

34. Ieronutti L, Chittaro L. Employing virtual humans for education and training in X3D/VRML worlds. Computers & Education 2007 Aug;49(1):93-109. [doi: 10.1016/j.compedu.2005.06.007]

35. Kleinsmith A, Rivera-Gutierrez D, Finney G, Cendan J, Lok B. Understanding Empathy Training with Virtual Patients. Comput Human Behav 2015 Nov 01;52:151-158 [FREE Full text] [doi: 10.1016/j.chb.2015.05.033] [Medline: 26166942]

36. Provoost S, Lau HM, Ruwaard J, Riper H. Embodied Conversational Agents in Clinical Psychology: A Scoping Review. J Med Internet Res 2017 May 09;19(5):e151 [FREE Full text] [doi: 10.2196/jmir.6553] [Medline: 28487267]

37. Laranjo L, Dunn A, Tong H, Kocaballi AB, Chen J, Bashir R, et al. Conversational agents in healthcare: a systematic review. J Am Med Inform Assoc 2018 Sep 01;25(9):1248-1258 [FREE Full text] [doi: 10.1093/jamia/ocy072] [Medline: 30010941]

38. Gaffney H, Mansell W, Tai S. Conversational Agents in the Treatment of Mental Health Problems: Mixed-Method Systematic Review. JMIR Ment Health 2019 Oct 18;6(10):e14166 [FREE Full text] [doi: 10.2196/14166] [Medline: 31628789]

39. Vaidyam AN, Wisniewski H, Halamka JD, Kashavan MS, Torous JB. Chatbots and Conversational Agents in Mental Health: A Review of the Psychiatric Landscape. Can J Psychiatry 2019 Jul 21;64(7):456-464 [FREE Full text] [doi: 10.1177/0706743719828977] [Medline: 30897957]

40. Kocaballi AB, Berkovsky S, Quiroz JC, Laranjo L, Tong HL, Rezazadegan D, et al. The Personalization of Conversational Agents in Health Care: Systematic Review. J Med Internet Res 2019 Nov 07;21(11):e15360 [FREE Full text] [doi: 10.2196/15360] [Medline: 31697237]

41. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009 Aug 18;151(4):264-9, W64. [doi: 10.7326/0003-4819-151-4-200908180-00135] [Medline: 19622511]

42. Leeflang MM, Deeks JJ, Takwoingi Y, Macaskill P. Cochrane diagnostic test accuracy reviews. Syst Rev 2013 Oct 07;2(1):82 [FREE Full text] [doi: 10.1186/2046-4053-2-82] [Medline: 24099098]

43. Macaskill P, Gatsonis C, Deeks J, Harbord R, Takwoingi Y. Chapter 10: analysing and presenting results. In: Deeks JJ, Bossuyt PM, Gatsonis C, editors. Cochrane handbook for systematic reviews of diagnostic test accuracy version 1.0. Hoboken, New Jersey: The Cochrane Collaboration; 2010.

44. Hopkins IM, Gower MW, Perez TA, Smith DS, Amthor FR, Casey Wimsatt F, et al. Avatar Assistant: Improving Social Skills in Students with an ASD Through a Computer-Based Intervention. J Autism Dev Disord 2011 Feb 2;41(11):1543-1555. [doi: 10.1007/s10803-011-1179-z] [Medline: 21287255]

45. EndNote®. URL: https://endnote.com/ [accessed 2020-07-08]

46.   Riva JJ, Malik KMP, Burnie SJ, Endicott AR, Busse JW. What is your research question? An introduction to the PICOT format for clinicians. J Can Chiropr Assoc 2012 Sep;56(3):167-171 [FREE Full text] [Medline: 22997465]

47.   Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011 Oct 18;343(oct18 2):d5928-d5928 [FREE Full text] [doi: 10.1136/bmj.d5928] [Medline: 22008217]

48.   Peters JL, Sutton AJ, Jones DR, Abrams KR, Rushton L. Comparison of two methods to detect publication bias in meta-analysis. JAMA 2006 Feb 08;295(6):676-680. [doi: 10.1001/jama.295.6.676] [Medline: 16467236]

49.   Simonsohn U, Nelson LD, Simmons JP. p-Curve and Effect Size: Correcting for Publication Bias Using Only Significant Results. Perspect Psychol Sci 2014 Nov 17;9(6):666-681. [doi: 10.1177/1745691614553988] [Medline: 26186117]

50.   Duval S, Tweedie R. Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 2000 Jun;56(2):455-463. [doi: 10.1111/j.0006-341x.2000.00455.x] [Medline: 10877304]

51.   meta: General Package for Meta-Analysis. 2012. URL: https://cran.r-project.org/web/packages/meta/index.html [accessed 2020-07-06]

52.   metafor: Meta-Analysis Package for R. URL: https://cran.r-project.org/web/packages/metafor/index.html [accessed 2020-07-07]

53.   Viechtbauer W. Conducting meta-analyses in R with the metafor package. J statistical software 2010 Aug 05;36(3):1-48.

54.   Lucas GM, Rizzo A, Gratch J, Scherer S, Stratou G, Boberg J, et al. Reporting Mental Health Symptoms: Breaking Down Barriers to Care with Virtual Human Interviewers. Front. Robot. AI 2017 Oct 12;4:1. [doi: 10.3389/frobt.2017.00051]

55.   DeVault D, Artstein R, Benn G, Dey T, Fast E, Gainer A, et al. SimSensei kiosk: a virtual human interviewer for healthcare decision support. 2014 Presented at: International conference on Autonomous Agents and Multi-Agent Systems; 2014; Paris, France p. 1061-1068. [doi: 10.5555/2615731.2617415]

56.   Zhou S, Bickmore T, Rubin A, Yeksigian C, Sawdy M, Simon S. User Gaze Behavior while Discussing Substance Use with a Virtual Agent. 2018 Presented at: International Conference on Intelligent Virtual Agents; 2018; Sydney NSW Australia p. 353-354. [doi: 10.1145/3267851.3267906]

57.   Ring L, Barry B, Totzke K, Bickmore T. Addressing Loneliness and Isolation in Older Adults: Proactive Affective Agents Provide Better Support. : IEEE; 2013 Presented at: 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction; 2013; Geneva, Switzerland. [doi: 10.1109/ACII.2013.17]

58.   Andrade AD, Anam R, Karanam C, Downey P, Ruiz JG. An overactive bladder online self-management program with embedded avatars: a randomized controlled trial of efficacy. Urology 2015 Mar;85(3):561-567. [doi: 10.1016/j.urology.2014.11.017] [Medline: 25586473]

59.   Magnani JW, Schlusser CL, Kimani E, Rollman BL, Paasche-Orlow MK, Bickmore TW. The Atrial Fibrillation Health Literacy Information Technology System: Pilot Assessment. JMIR Cardio 2017 Dec 12;1(2):e7 [FREE Full text] [doi: 10.2196/cardio.8543] [Medline: 29473644]

60.   Bickmore T, Kimani E, Trinh H, Pusateri A, Paasche-Orlow M, Magnani J. Managing Chronic Conditions with a Smartphone-based Conversational Virtual Agent. 2018 Presented at: International Conference on Intelligent Virtual Agents; 2018; Sydney NSW Australia p. 119-124. [doi: 10.1145/3267851.3267908]

61.   Kimani E, Bickmore T, Trinh H, Ring L, Paasche-Orlow M, Magnani J. A Smartphone-Based Virtual Agent for Atrial Fibrillation Education and Counseling. 2016 Presented at: International Conference on Intelligent Virtual Agents; 2016; Los Angeles, USA p. 120-127. [doi: 10.1007/978-3-319-47665-0_11]

62.   Shamekhi A, Trinh H, Bickmore T, DeAngelis T, Ellis T, Houlihan B, et al. A Virtual Self-care Coach for Individuals with Spinal Cord Injury. 2016 Presented at: 18th International ACM SIGACCESS Conference on Computers and Accessibility - ASSETS '16; 2016; Reno Nevada USA p. 327-328. [doi: 10.1145/2982142.2982199]

63.   Bickmore TW, Silliman RA, Nelson K, Cheng DM, Winter M, Henault L, et al. A randomized controlled trial of an automated exercise coach for older adults. J Am Geriatr Soc 2013 Oct 03;61(10):1676-1683. [doi: 10.1111/jgs.12449] [Medline: 24001030]

64.   Bickmore TW, Caruso L, Clough-Gorr K, Heeren T. 'It's just like you talk to a friend' relational agents for older adults. Interacting with Computers 2005 Dec;17(6):711-735. [doi: 10.1016/j.intcom.2005.09.002]

65.   Bickmore TW, Pfeifer LM, Byron D, Forsythe S, Henault LE, Jack BW, et al. Usability of conversational agents by patients with inadequate health literacy: evidence from two clinical trials. J Health Commun 2010 Aug 31;15 Suppl 2(sup2):197-210. [doi: 10.1080/10810730.2010.499991] [Medline: 20845204]

66.   Ellis T, Latham NK, DeAngelis TR, Thomas CA, Saint-Hilaire M, Bickmore TW. Feasibility of a Virtual Exercise Coach to Promote Walking in Community-Dwelling Persons with Parkinson Disease. American Journal of Physical Medicine & Rehabilitation 2013;92(6):472-485. [doi: 10.1097/phm.0b013e31828cd466]

67.   King AC, Bickmore TW, Campero MI, Pruitt LA, Yin JL. Employing virtual advisors in preventive care for underserved communities: results from the COMPASS study. J Health Commun 2013 Dec;18(12):1449-1464 [FREE Full text] [doi: 10.1080/10810730.2013.798374] [Medline: 23941610]

68.   Friederichs S, Bolman C, Oenema A, Guyaux J, Lechner L. Motivational interviewing in a Web-based physical activity intervention with an avatar: randomized controlled trial. J Med Internet Res 2014 Feb 13;16(2):e48 [FREE Full text] [doi: 10.2196/jmir.2974] [Medline: 24550153]

69. Bickmore TW, Schulman D, Sidner C. Automated interventions for multiple health behaviors using conversational agents. Patient Educ Couns 2013 Aug;92(2):142-148 [FREE Full text] [doi: 10.1016/j.pec.2013.05.011] [Medline: 23763983]

70. Bickmore TW, Mauer D, Brown T. Context Awareness in a Handheld Exercise Agent. Pervasive Mob Comput 2009 Jun 01;5(3):226-235 [FREE Full text] [doi: 10.1016/j.pmcj.2008.05.004] [Medline: 20161031]

71. Bickmore TW, Puskar K, Schlenk EA, Pfeifer LM, Sereika SM. Maintaining reality: Relational agents for antipsychotic medication adherence. Interacting with Computers 2010 Jul;22(4):276-288. [doi: 10.1016/j.intcom.2010.02.001]

72. Burton C, Szentagotai Tatar A, McKinstry B, Matheson C, Matu S, Moldovan R, et al. Pilot randomised controlled trial of Help4Mood, an embodied virtual agent-based system to support treatment of depression. J Telemed Telecare 2016 Jul 09;22(6):348-355. [doi: 10.1177/1357633x15609793]

73. Ring L, Bickmore T, Pedrelli P. An affectively aware virtual therapist for depression counseling. 2016 Presented at: SIGCHI Conference on Human Factors in Computing Systems (CHI) workshop on Computing and Mental Health; 2016; Denver Colorado USA.

74. Han K, Heo J, Seo S, Hong M, Lee JS, Shin YS, et al. The effect of simulated auditory hallucinations on daily activities in schizophrenia patients. Psychopathology 2012 Jul 31;45(6):352-360 [FREE Full text] [doi: 10.1159/000337264] [Medline: 22854179]

75. Gardiner PM, McCue KD, Negash LM, Cheng T, White LF, Yinusa-Nyahkoon L, et al. Engaging women with an embodied conversational agent to deliver mindfulness and lifestyle recommendations: A feasibility randomized control trial. Patient Educ Couns 2017 Sep;100(9):1720-1729 [FREE Full text] [doi: 10.1016/j.pec.2017.04.015] [Medline: 28495391]

76. Shamekhi A, Bickmore T, Lestoquoy A, Gardiner P. Augmenting Group Medical Visits with Conversational Agents for Stress Management Behavior Change. Persuasive Technology: Development and Implementation of Personalized Technologies to Change Attitudes and Behaviorspringer International Publishing; 2017 Presented at: International Conference on Persuasive Technology; 2017; Amsterdam, The Netherlands p. 55-67. [doi: 10.1007/978-3-319-55134-0_5]

77. Shamekhi A, Bickmore T, Lestoquoy A, Negash L, Gardiner P. Blissful Agents: Adjuncts to Group Medical Visits for Chronic Pain and Depression. Blissful Agents: Adjuncts to Group Medical Visits for Chronic Pain and Depression. In. Intelligent Virtual Agentspringer International Publishing; 2016 Presented at: International Conference on Intelligent Virtual Agents; 2016; Los Angeles, USA p. 433-437. [doi: 10.1007/978-3-319-47665-0_49]

78. Park K, Ku J, Choi S, Jang H, Park J, Kim SI, et al. A virtual reality application in role-plays of social skills training for schizophrenia: a randomized, controlled trial. Psychiatry Res 2011 Sep 30;189(2):166-172. [doi: 10.1016/j.psychres.2011.04.003] [Medline: 21529970]

79. Mountford VA, Tchanturia K, Valmaggia L. "What Are You Thinking When You Look at Me?" A Pilot Study of the Use of Virtual Reality in Body Image. Cyberpsychol Behav Soc Netw 2016 Feb;19(2):93-99. [doi: 10.1089/cyber.2015.0169] [Medline: 26378613]

80. Utami D, Bickmore T, Nikolopoulou A, Paasche-Orlow M. Talk About Death: End of Life Planning with a Virtual Agent. 2017 Presented at: International Conference on Intelligent Virtual Agents; 2017; Stockholm, Sweden p. 441-450. [doi: 10.1007/978-3-319-67401-8_55]

81. Bickmore TW, Utami D, Matsuyama R, Paasche-Orlow MK. Improving Access to Online Health Information With Conversational Agents: A Randomized Controlled Experiment. J Med Internet Res 2016 Jan 04;18(1):e1 [FREE Full text] [doi: 10.2196/jmir.5239] [Medline: 26728964]

82. Utami D, Barry B, Bickmore T, Paasche-Orlow M. A conversational agent-based clinical trial search engine. 2013 Presented at: Annual Symposium on Human-Computer Interaction and Information Retrieval; 2013; Vancouver, BC, Canada.

83. Bickmore TW, Pfeifer LM, Paasche-Orlow MK. Using computer agents to explain medical documents to patients with low health literacy. Patient Educ Couns 2009 Jun;75(3):315-320 [FREE Full text] [doi: 10.1016/j.pec.2009.02.007] [Medline: 19297116]

84. Bickmore T, Pfeifer L, Jack B. Taking the time to care: empowering low health literacy hospital patients with virtual nurse agents. 2009 Presented at: SIGCHI conference on human factors in computing systems; 2009; Boston MA USA p. 1265-1274. [doi: 10.1145/1518701.1518891]

85. Bickmore T, Pfeifer L, Paasche-Orlow M. Health Document Explanation by Virtual Agents. 2007 Presented at: International Workshop on Intelligent Virtual Agents; 2007; Paris, France p. 183-196. [doi: 10.1007/978-3-540-74997-4_18]

86. Zhou S, Bickmore T, Paasche-Orlow M, Jack B. Agent-User Concordance and Satisfaction with a Virtual Hospital Discharge Nurse. 2014 Presented at: International Conference on Intelligent Virtual Agents; 2014; Boston, MA, USA p. 528-541. [doi: 10.1007/978-3-319-09767-1_63]

87. Schulman D, Bickmore T. Persuading users through counseling dialogue with a conversational agent. 2009 Presented at: International Conference on Persuasive Technology; 2009; Claremont California USA p. 1-8. [doi: 10.1145/1541948.1541983]

88. Edwards RA, Bickmore T, Jenkins L, Foley M, Manjourides J. Use of an interactive computer agent to support breastfeeding. Matern Child Health J 2013 Dec 18;17(10):1961-1968. [doi: 10.1007/s10995-013-1222-0] [Medline: 23329167]

89. Zhang Z, Bickmore T, Mainello K, Mueller M, Foley M, Jenkins L, et al. Maintaining Continuity in Longitudinal, Multi-method Health Interventions Using Virtual Agents: The Case of Breastfeeding Promotion. 2014 Presented at: International Conference on Intelligent Virtual Agents; 2014; Boston, MA, USA p. 504-513. [doi: 10.1007/978-3-319-09767-1_61]

90. Shi L, Bickmore T, Edwards R. A Feminist Virtual Agent for Breastfeeding Promotion. 2015 Presented at: International Conference on Intelligent Virtual Agents; 2015; Delft, The Netherlands p. 461-470. [doi: 10.1007/978-3-319-21996-7_49]

91. Manuvinakurike R, Velicer WF, Bickmore TW. Automated indexing of Internet stories for health behavior change: weight loss attitude pilot study. J Med Internet Res 2014 Dec 09;16(12):e285 [FREE Full text] [doi: 10.2196/jmir.3702] [Medline: 25491243]

92. Ruiz JG, Andrade AD, Anam R, Lisigurski M, Karanam C, Sharit J. Computer-based programmed instruction did not improve the knowledge retention of medication instructions of individuals with type 2 diabetes mellitus. Diabetes Educ 2014 Oct 29;40(1):77-88. [doi: 10.1177/0145721713508825] [Medline: 24168837]

93. Zhang Z, Bickmore T. Medical Shared Decision Making with a Virtual Agent. 2018 Presented at: International Conference on Intelligent Virtual Agents; 2018; Sydney NSW Australia p. 113-118. [doi: 10.1145/3267851.3267883]

94. Cassidy S, Stenger B, Van Dongen L, Yanagisawa K, Anderson R, Wan V, et al. Expressive visual text-to-speech as an assistive technology for individuals with autism spectrum conditions. Comput Vis Image Underst 2016 Jul;148:193-200 [FREE Full text] [doi: 10.1016/j.cviu.2015.08.011] [Medline: 27375348]

95. Marcos-Pablos S, González-Pablos E, Martín-Lorenzo C, Flores LA, Gómez-García-Bermejo J, Zalama E. Corrigendum: Virtual Avatar for Emotion Recognition in Patients with Schizophrenia: A Pilot Study. Front Hum Neurosci 2016 Nov 02;10:554 [FREE Full text] [doi: 10.3389/fnhum.2016.00554] [Medline: 27807413]

96. Kim K, Rosenthal MZ, Gwaltney M, Jarrold W, Hatt N, McIntyre N, et al. A Virtual Joy-Stick Study of Emotional Responses and Social Motivation in Children with Autism Spectrum Disorder. J Autism Dev Disord 2015 Dec 17;45(12):3891-3899 [FREE Full text] [doi: 10.1007/s10803-014-2036-7] [Medline: 24435405]

97. Gardiner P, Hempstead MB, Ring L, Bickmore T, Yinusa-Nyahkoon L, Tran H, et al. Reaching Women through Health Information Technology: The Gabby Preconception Care System. Am J Health Promot 2013 Jan 02;27(3_suppl):eS11-eS20. [doi: 10.4278/ajhp.1200113-quan-18]

98. Hesse K, Schroeder PA, Scheeff J, Klingberg S, Plewnia C. Experimental variation of social stress in virtual reality - Feasibility and first results in patients with psychotic disorders. J Behav Ther Exp Psychiatry 2017 Sep;56:129-136. [doi: 10.1016/j.jbtep.2016.11.006] [Medline: 27939053]

99. Jarrold W, Mundy P, Gwaltney M, Bailenson J, Hatt N, McIntyre N, et al. Social attention in a virtual public speaking task in higher functioning children with autism. Autism Res 2013 Oct 20;6(5):393-410 [FREE Full text] [doi: 10.1002/aur.1302] [Medline: 23696132]

100. Shamekhi A, Bickmore T. Breathe Deep: A Breath-Sensitive Interactive Meditation Coach. 2018 Presented at: International Conference on Pervasive Computing Technologies for Healthcare; 2018; New York NY USA. [doi: 10.1145/3240925.3240940]

101. Olafsson S, O'Leary T, Bickmore T. Coerced Change-talk with Conversational Agents Promotes Confidence in Behavior Change. 2019 Presented at: International Conference on Pervasive Computing Technologies for Healthcare; 2019; Trento, Italy p. 31-40. [doi: 10.1145/3329189.3329202]

102. Zhang Z, Bickmore TW, Paasche-Orlow MK. Perceived organizational affiliation and its effects on patient trust: Role modeling with embodied conversational agents. Patient Educ Couns 2017 Sep;100(9):1730-1737. [doi: 10.1016/j.pec.2017.03.017] [Medline: 28381330]

103. Wieser MJ, Pauli P, Grosseibl M, Molzow I, Mühlberger A. Virtual social interactions in social anxiety--the impact of sex, gaze, and interpersonal distance. Cyberpsychol Behav Soc Netw 2010 Oct;13(5):547-554. [doi: 10.1089/cyber.2009.0432] [Medline: 20950179]

104. Gunn C, Maschke A, Bickmore T, Kennedy M, Hopkins MF, Fishman MD, et al. Acceptability of an Interactive Computer-Animated Agent to Promote Patient-Provider Communication About Breast Density: a Mixed Method Pilot Study. J Gen Intern Med 2020 Apr 9;35(4):1069-1077. [doi: 10.1007/s11606-019-05622-2] [Medline: 31919723]

105. Dworkin MS, Lee S, Chakraborty A, Monahan C, Hightow-Weidman L, Garofalo R, et al. Acceptability, Feasibility, and Preliminary Efficacy of a Theory-Based Relational Embodied Conversational Agent Mobile Phone Intervention to Promote HIV Medication Adherence in Young HIV-Positive African American MSM. AIDS Educ Prev 2019 Feb;31(1):17-37. [doi: 10.1521/aeap.2019.31.1.17] [Medline: 30742481]

106. Bickmore T. Relational agents: Effecting change through human-computer relationships. Massachusetts, USA: Doctoral dissertation, Massachusetts Institute of Technology; 2003.

107. Cassell J, Vilhjálmsson H, Bickmore T. BEAT: the Behavior Expression Animation Toolkit. 2001 Presented at: International Conference on Computer Graphics and Interactive Techniques, SIGGRAPH; 2001; Los Angeles, USA p. 477-486. [doi: 10.1007/978-3-662-08373-4_8]

108. Woods WA. Transition network grammars for natural language analysis. Commun. ACM 1970;13(10):591-606. [doi: 10.1145/355598.362773]

109. Bodenheimer T, Wagner EH, Grumbach K. Improving primary care for patients with chronic illness: the chronic care model, Part 2. JAMA 2002 Oct 16;288(15):1909-1914. [doi: 10.1001/jama.288.15.1909] [Medline: 12377092]

110. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997 Aug 26;12(1):38-48. [doi: 10.4278/0890-1171-12.1.38] [Medline: 10170434]

111. Rollnick S, Miller WR. What is Motivational Interviewing? Behav. Cogn. Psychother 2009 Jun 16;23(4):325-334. [doi: 10.1017/s135246580001643x]

112. Soller A, Martínez A, Jermann P, Muehlenbrock M. From mirroring to guiding: A review of state of the art technology for supporting collaborative learning. Intl J Artificial Intelligence in Education 2005;15(4):261-290.

113. Stacey D, Murray MA, Légaré F, Sandy D, Menard P, O'Connor A. Decision coaching to support shared decision making: a framework, evidence, and implications for nursing practice, education, and policy. Worldviews Evid Based Nurs 2008 Mar;5(1):25-35. [doi: 10.1111/j.1741-6787.2007.00108.x] [Medline: 18266768]

114. Elwyn G, Frosch D, Thomson R, Joseph-Williams N, Lloyd A, Kinnersley P, et al. Shared decision making: a model for clinical practice. J Gen Intern Med 2012 Oct 23;27(10):1361-1367 [FREE Full text] [doi: 10.1007/s11606-012-2077-6] [Medline: 22618581]

115. Viechtbauer W, Cheung MW. Outlier and influence diagnostics for meta-analysis. Res Synth Methods 2010 Apr 04;1(2):112-125. [doi: 10.1002/jrsm.11] [Medline: 26061377]

116. Baujat B, Mahé C, Pignon J, Hill C. A graphical method for exploring heterogeneity in meta-analyses: application to a meta-analysis of 65 trials. Stat Med 2002 Sep 30;21(18):2641-2652. [doi: 10.1002/sim.1221] [Medline: 12228882]

117. Aguinis H, Gottfredson RK, Joo H. Best-Practice Recommendations for Defining, Identifying, and Handling Outliers. Organizational Research Methods 2013 Jan 14;16(2):270-301. [doi: 10.1177/1094428112470848]

118. Gratch J, DeVault D, Lucas G. The Benefits of Virtual Humans for Teaching Negotiation. 2016 Presented at: International Conference on Intelligent Virtual Agents; 2016; Los Angeles, USA p. 283-294. [doi: 10.1007/978-3-319-47665-0_25]

119. Carnell S, Lok B, James M, Su J. Predicting Student Success in Communication Skills Learning Scenarios with Virtual Humans. 2019 Presented at: International Conference on Learning Analytics & Knowledge; 2019; Tempe, AZ, USA p. 436-440. [doi: 10.1145/3303772.3303828]

120. Feng Z, González VA, Amor R, Lovreglio R, Cabrera-Guerrero G. Immersive virtual reality serious games for evacuation training and research: A systematic literature review. Computers & Education 2018 Dec;127:252-266. [doi: 10.1016/j.compedu.2018.09.002]

121. Bickmore TW, Trinh H, Olafsson S, O'Leary TK, Asadi R, Rickles NM, et al. Patient and Consumer Safety Risks When Using Conversational Assistants for Medical Information: An Observational Study of Siri, Alexa, and Google Assistant. J Med Internet Res 2018 Sep 04;20(9):e11510 [FREE Full text] [doi: 10.2196/11510] [Medline: 30181110]

122. Yellowlees PM, Holloway KM, Parish MB. Therapy in virtual environments--clinical and ethical issues. Telemed J E Health 2012 Sep;18(7):558-564. [doi: 10.1089/tmj.2011.0195] [Medline: 22823138]

123. Witt KJ, Oliver M, McNichols C. Counseling via Avatar: Professional Practice in Virtual Worlds. Int J Adv Counselling 2016 Jul 2;38(3):218-236. [doi: 10.1007/s10447-016-9269-4]

124. Parmar D, Olafsson S, Utami D, Bickmore T. Looking the part: The effect of attire and setting on perceptions of a virtual health counselor. 2018 Presented at: International Conference on Intelligent Virtual Agents; 2018; Sydney, NSW, Australia p. 301-306. [doi: 10.1145/3267851.3267915]

125. Ring L, Utami D, Bickmore T. The right agent for the job? 2014 Presented at: International Conference on Intelligent Virtual Agents; 2014; Boston, MA, USA p. 374-384. [doi: 10.1007/978-3-319-09767-1_49]

126. Meah L, Moore R. The Uncanny Valley: A Focus on Misaligned Cues. 2014 Presented at: International Conference on Social Robotics; 2014; Sydney, NSW, Australia p. 256-265. [doi: 10.1007/978-3-319-11973-1_26]

Abbreviations

3D: 3-dimensional.
a-PDHA: anonymized post-deployment health assessment.
ACT: physical activity.
ASD: autism spectrum disorders.
ATN: augmented transition network.
BDI-2: Beck Depression Inventory-II.
BEAT: Behavior Expression Animation Toolkit.
BICEP: brief informed consent evaluation protocol.
DAS-SF2: Dysfunctional Attitude Scale-Short Form 2.
DIET: fruit and vegetable consumption.
EQ−5D−5L VAS: 5-level version of the EuroQol 5D visual analogue scale.
FVS: NIH/NCI Fruit and Vegetable Scan.
HMD: head-mounted display.
HRQOL: health-related quality of life.
mHealth: mobile health.
MI: motivational interviewing.
OABq: overactive bladder questionnaire.
PDHA: post-deployment health assessment.

PTSD: post-traumatic stress disorder.
QIDS-SR: Quality of Life Enjoyment and Satisfaction Questionnaire-Short Form.
RCT: randomized controlled trial.
SBS: social behavior scales.
SMD: standardized mean difference.
SVH: social virtual human.
VH: virtual human.
VR: virtual reality.

XSL•FO
RenderX