

Article

Social Responses to Virtual Humans: The Effect of Human-Like Characteristics

Sung Park ^{*,†}  and Richard Catrambone 

School of Psychology, Georgia Institute of Technology, Atlanta, GA 30332, USA;
richard.catrambone@psych.gatech.edu

* Correspondence: spica7601@gmail.com

† Sung Park is now at School of Design, Savannah College of Art and Design.

Abstract: As a virtual human is provided with more human-like characteristics, will it elicit stronger social responses from people? Two experiments were conducted to address these questions. The first experiment investigated whether virtual humans can evoke a social facilitation response and how strong that response is when people are given different cognitive tasks that vary in difficulty. The second experiment investigated whether people apply politeness norms to virtual humans. Participants were tutored either by a human tutor or a virtual human tutor that varied in features and then evaluated the tutor's performance. Results indicate that virtual humans can produce social facilitation not only with facial appearance but also with voice. In addition, performance in the presence of voice synced facial appearance seems to elicit stronger social facilitation than in the presence of voice only or face only. Similar findings were observed with the politeness norm experiment. Participants who evaluated their tutor directly reported the tutor's performance more favorably than participants who evaluated their tutor indirectly. This valence toward the voice synced facial appearance had no statistical difference compared to the valence toward the human tutor condition. The results suggest that designers of virtual humans should be mindful about the social nature of virtual humans.



Citation: Park, S.; Catrambone, R. Social Responses to Virtual Humans: The Effect of Human-Like Characteristics. *Appl. Sci.* **2021**, *11*, 7214. <https://doi.org/10.3390/app11167214>

Academic Editor: Andrea Prati

Received: 20 July 2021

Accepted: 2 August 2021

Published: 5 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: virtual human; virtual character; embodied conversational agent; avatar; anthropomorphism; social facilitation; politeness norm; social response

1. Introduction

Interest in virtual humans or embodied conversational agents is growing in the realm of human computer interaction (HCI). Improvements in technology have facilitated the use of virtual humans in various applications, such as entertainment [1,2], engineering [3,4], medical practice [5,6], and the military [7,8]. Specifically, the VR industry is emerging at a rapid rate, with widespread recognition of the use of virtual humans in gaming [9], healthcare [10], and education [11]. Virtual humans are embodied as agents controlled by an AI or an avatar, an alter-ego of a user [12]. Because the common denominator of a virtual human is human-like characteristics, virtual humans do not necessarily need a visual form. The public has experienced the release of voice recognition assistants such as Amazon's Echo (2014), Microsoft's Cortana (2014), and Apple's Siri (2011). Samsung Electronics has recently released Samsung Girl (2021), adding personalized facial features, gaining media attention because of the unique differentiator from their competitor's virtual personas limited to voice.

The work of Nass and other researchers [13–15] suggests that there is a striking similarity between how humans interact with one another and how a human and a computer interact. The heart of this similarity is the socialness of the interaction; as Nass, Steuer, and Tauber stated, “computers are social actors” [16].

Our research is directed towards gaining a deeper understanding of the social dimension of the interaction between humans and virtual humans. Specifically, we are interested

in whether a virtual human elicits stronger social responses as it is provided with more human-like characteristics.

We reviewed the current literature on human-computer interaction, human-robot interaction, and ordinary personology to understand the cognitive process behind the social responses to virtual humans. We established a qualitative framework of how humans respond socially to virtual humans; the framework proposes that as long as a human-like characteristic is exhibited, there is some possibility that a virtual human will be categorized as a human and elicit a social response. The number of human-like characteristics may impact the likelihood with which a virtual human is categorized as human. To test this hypothesis, we conducted two experiments: (1) a social facilitation experiment and (2) a politeness norm experiment. Details on methods and results will follow in addition of the discussion to their generalizability and limitations.

1.1. Social Responses to Virtual Humans

Being social refers to people thinking about, influencing, and relating to one another [17]. Being social constitutes many behaviors (e.g., perception of others and judgment of others) and psychological constructs (e.g., belief, persuasion, attraction). In the human computer interaction literature, responding socially to virtual humans means that people exhibit behavioral responses to virtual humans as if they are humans. Based partly on the work of Nass and his colleagues [15,16] (for a review, see [18]), there is general agreement in the research community that people do respond socially to computers. Nass's research method is essentially a Turing test [19] but assesses non-human entities' social capability rather than a machine's ability to demonstrate intelligence. His procedure is to take a well-established study from social psychology, replace the person in the study with a computer, and assess whether the results of the study are equivalent to the original study with a person, that is, whether participants behave in the same way when interacting with a computer as they would have with a human. In their studies, participants applied many social rules (e.g., ethnicity stereotype, emotional consistency) to their behaviors towards computers that are normally observed when interacting with humans. Although most follow-up research investigated the presence or absence of computer-as-social-actors, research attempts to understand what contributes to such social effects are scarce. To gain a deeper insight, we will first zoom out and review humans' social responses to humans.

1.2. How Humans Respond Socially to Humans

"Ordinary Personology" is the study of how an individual comes to know about another individual's temporary states (emotions, intentions, and desires) and enduring dispositions (beliefs, traits, and abilities). Gilbert [20] suggested a framework which attempted to incorporate attribution theories and social cognitive theories. Consider the following example used by Gilbert to illustrate the framework. An actor who has a particular appearance (e.g., wearing a crucifix or a Mohawk) that allows him to be classified as a member of a category (e.g., a priest or a punk) will then allow the observer to draw inferences about the actor's dispositions (he is religious or rebellious). In addition to the discrete items of information gathered from the disposition process (e.g., he should be religious because he is a priest), we also gather information from observation to find a unifying explanation for those features. The key to understanding the social responses to virtual humans may lie with this categorization and inferential process.

Humans inevitably categorize objects and people in their world [21]. While Gilbert's aforementioned example assumes that the actor is a human and in fact involves a property of a person, the framework can be equally applied to how individuals categorize any entity, including a virtual human, as a human. Categorization and determining disposition are rapid and automatic due to their resource saving purposes [22]. Automaticity is the idea that sufficient practice with tasks maps stimuli and responses consistently and therefore produces performance that is autonomous, involuntary, unconscious, and undemanding of cognitive resources [23,24] (for a review, see [25]). The automatic nature of social responses

is the result of constant pairing of stimuli and social behaviors and therefore inherits all of the aforementioned characteristics.

In summary, the grounds for social responses are established when the human readily perceives the entity's human-like characteristics. Such human-like cues precipitate "human" categorization and disposition processes, which are unconscious and automatic. Virtual humans are unique among other anthropomorphic entities due to inherent and obtrusive human-like characteristics and so merit further examination.

1.3. How Humans Respond Socially to Virtual Humans

Three early theories attempted to understand the social response to virtual humans. Kiesler and Sproull [26] claimed that the social response was due to the novelty of the situation and would disappear once the person became knowledgeable about the interaction and technology. More elaborate models include the aforementioned computer-as-social-actors by Nass and his colleagues [13–15] and the threshold model of social influence by Blascovich et al. [27] (for a complete review see [28]). The essence of Nass' claim is that people unconsciously respond socially to computers, which was supported with ample empirical data of participants' categorical denial that they did not think there were programmers behind computers.

Blascovich et al. [27], in their Threshold Model of Social Influence, claim that a social response is a function of two factors: perception of agency and behavior realism. Agency is high when the human knows that the virtual human is a representation of a human being; that is, in a form of an avatar. Conversely, the agency is low when the human knows that the virtual human is a representation of a computer program. Both are on continuous dimensions, from low behavioral realism and low agency (agent) to high behavioral realism and high agency (avatar). When the two factors are crossed and the threshold is met, social responses are evoked. This means that when the agency is high, behavioral realism does not necessarily have to be high to elicit social responses and vice versa.

Putten, Kramer, Gratch and Kang [28] took the agent-avatar paradigm further and varied the participant's belief (avatar versus agent) and varied the dynamic listening behavior (realism) of the virtual human on a self-disclosure task. They found that higher behavior realism, which was manipulated as providing or not providing feedback after the disclosure, affected the evaluation of the virtual human whereas the belief did not. This finding may deny the effect of agency (agent versus avatar) in Blascovich's theory but supports Nass' speculation that the more computers present human-like characteristics, the more likely they will elicit social behavior [18] (p. 97). Similar findings were reported with robots [29].

Fox et al. [30] claimed that the reason studies had inconsistent results is because social responses are moderated with the level of immersion (desktop vs. fully immersive), dependent variable type (subjective vs. objective), task type (competitive vs. cooperative vs. neutral), and the control of the representation (human vs. computer).

Indeed, the presence and strength of social response seem to be contingent on the nature of task and how it was measured. Given that there are so many applications and tasks a virtual human can do, it may be difficult to find a main effect on the strength of social responses applicable to all domains. One way to approach the problem is to focus on the mere presence of a virtual human and the additive effect of the human-like characteristics and simultaneously minimize the interaction with the virtual human. This may sacrifice the ecological validity of the findings, but would provide clarity in understanding the cognitive mechanics of social responses and provide a useful research framework for future investigation with more complex interactions and tasks.

1.4. Virtual Human Characteristics

What are the particular characteristics of virtual humans that encourage social responses? Investigating human-like cues naturally leads us to think what constitutes a human. Some characteristics that come to mind are the ability to talk, listen, express emo-

tion, and look like a human. People recognize entities to have minds when they show signs of emotion and intelligence [31]. In this paper we consider the physical characteristics of virtual humans, including facial and voice embodiment, and the cognitive characteristics of virtual humans, including expression of emotion and personality. These characteristics are selected based on prior research. Specifically, empirical studies are included only if there are: (1) a clear social response or social behavior involved with a virtual human, and (2) an observable human-like cue embodied in a virtual human. Note that we included many empirical studies by Nass and his colleagues because while he used “computer” as a construct for his manipulation, his experimental materials fall under our definition of virtual humans which is defined as non-human entities that include human-like characteristics.

1.4.1. Facial Embodiment

Facial embodiment not only includes the embodiment of a human-like face but also non-verbal facial behaviors such as smiles, raising eyebrows, head movements, and gaze. Yee et al. [32] reviewed empirical studies (12 studies with performance measures and 22 studies with subjective measures) that compared interfaces with facial embodiment to interfaces without facial embodiment and found that the presence of a facial representation produced more positive interactions than the absence of a facial representation. This effect was found both in performance measures and subjective measures.

While facial embodiment in a few studies [33,34] was confounded with other human-like characteristics, facial embodiment in the other studies [35,36] was the sole characteristic embodied in a virtual human. Some studies, mostly Nass’s studies, investigated social rules that had been well known in social psychology and tested participants interacting with virtual humans to see whether the social rules still held. Other studies [36] tested participants in a “human” condition as a control group.

In experiments with a virtual human with facial embodiment, the social rule was observable or the virtual humans were as effective as humans in simulating a social phenomenon such as ethnicity stereotyping [35], persuasion [34], punitive ostracism [33], and social facilitation [37].

Clearly, facial embodiment or facial expression is a human-like characteristic that elicits social responses. This is understandable because the human face is a powerful signal of human identity [38]. Infants exhibit a preference for face-like patterns over other patterns [39]; they begin to differentiate facial features by the age of two months [40]. Faces induce appropriate social behaviors in particular situations. For example, covering people’s faces with masks can produce inappropriate behaviors [41] that fail to activate human categorization.

1.4.2. Voice Embodiment

While some studies [42,43] investigated the effect of voice exclusively, other studies investigated characteristics such as emotion or personality manifested in voice [44,45]. In experiments with a voice embodied in a virtual human, the social rule was observable or the virtual humans were as effective as humans in engendering a social phenomenon. Voice is probably the most studied human-like characteristics of virtual humans that produce social responses of consistency-attraction [44], emotional consistency [45], gender stereotype [16,46], notion of “self” and “other” [16], persuasiveness [34,42,43], politeness norm [14,16], and similarity attraction [44].

In short, a human-like cue via voice is sufficient to encourage users to apply social rules. Individuals unconsciously applied social rules despite consciously knowing that virtual humans are not people [47], and despite being presented with a voice that was not human (e.g., synthesized voice; [45]). Eyssel et al. [48] revealed that the same gender, between the virtual human and the participant, voice increased anthropomorphism more than different gender voice but only with a recorded human voice as opposed to a synthesized robotic one.

1.4.3. Emotion

Emotion plays a powerful role in people's lives. Emotion impacts our beliefs, informs our decision-making, and affects our social relationships. Humans have developed a range of behaviors that express emotional information as well as an ability to recognize other's emotions. There is no universally accepted definition of emotion, nor is there a general agreement as to which phenomena are accepted as the manifestation of emotion [49–51]. Some theories support a set of distinct emotions with neurological correlates [52] whereas others argue that emotions are epiphenomenal, simply reflecting the interaction of underlying processes [53]. The relationship between emotion and other psychological constructs such as feeling, mood, and personality is also unclear [49]. Izard claimed [54], based on current data, emotion cannot be defined as a unitary concept.

Nevertheless, the emergence of affective computing [55] provided tremendous insights to designing effective systems and services. Advances in creating such systems include the implementation of “affective loop”, the stages of recognition, understanding, and expression of emotion, not only in Human-Computer Interaction [56] but also in Human-Robot Interaction [57].

Clearly, emotion plays a major role in effective communication with virtual humans [45] and their believability [58]. Significant research suggests that the believability of emotional expression is related to the context where it occurs and should be expressed in relation to identifiable stimuli [49,59,60].

Social phenomenon such as emotional consistency [45] and punitive ostracism [33] are observed with virtual humans where emotion was manifested in voice and facial expression, respectively.

Emotion is also closely related to time. In our framework, emotion plays a stronger role in long-term interactions due to the history of interaction between users and virtual humans. For example, a negative emotional experience, such as frustration at an improper emotional expression, due to an asocial characteristic of a virtual human might influence subsequent interactions. However, the rapport built between the virtual human and the human in a therapeutic session seems to persist throughout time [61]. The participant's impression of the virtual human (empathic or neutral) in the first session where participants expressed their emotional feelings did not change even when the virtual human changed (empathic or neutral).

1.4.4. Personality

Personality is defined by behavioral consistency within a class of situations that is defined by the person [62]. A few studies employed personality as a human-like characteristic to investigate a particular social response such as consistency-attraction [44,63], gain-loss theory [64], self-service bias [65], and similarity-attraction [44,66].

In the present context, of interest is whether people respond to virtual humans as if virtual humans have a personality. A number of studies suggest that there are at least three factors that play a role in the social interaction with virtual humans involving personality, including characteristics of language [64,66], character appearance [67], and character behavior [15,68]. The human's personality also plays a role [69]. Specifically, the research showed that if the personalities are similar between the human and the virtual human, the human recognizes the virtual human's personality better than when the personalities are dissimilar.

1.5. Framework of Social Responses and Hypotheses

Four characteristics are identified through prior research as human-like characteristics: personality, emotion, facial embodiment, and voice. While facial embodiment and voice implementation can each directly elicit social responses [47], personality and emotion presumably work indirectly through facial expression, voice, or specific word choices.

There are other variables such as intelligence or autonomy that need further investigation. King and Ohya [70] found that human forms presented on a computer screen

are assessed as more intelligent than simple geometric forms. Paiva et al. [71] suggested that autonomy is one of the key features for the achievement of believability. Nevertheless, none of the studies have provided empirical evidence that virtual humans embodied with such constructs elicit social responses.

Figure 1 presents the complete framework. Human-like characteristics of virtual humans act as a cue that leads a person to place the agent into the category “human” and thus, elicits social responses. Social responding is so ingrained through practice that a human-like cue will set off social responses (i.e., automatic over-reliance), which are automatic and therefore generally unconscious [25]. Specifically, a single cue might activate the human category and block other cues (i.e., asocial cues) that would activate the virtual human category.

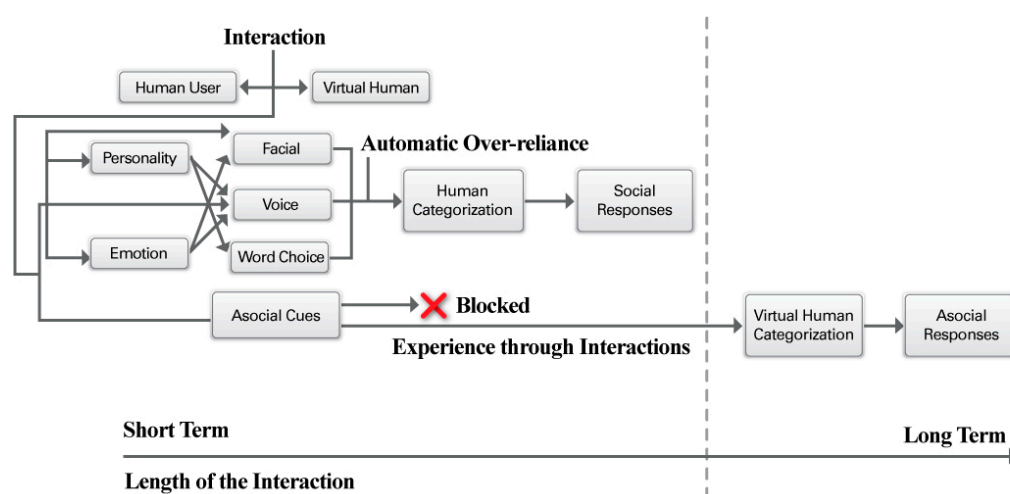


Figure 1. The framework of how humans respond socially to virtual humans.

However, many factors such as knowledge of virtual humans, motivation, accessibility, and the length of interaction can play a contributing role in ascribing the cue to the virtual human category rather than the human category. The more time spent with a virtual human, the greater chance there is to identify the asocial nature of a virtual human. That is, one will deliberately map the stimulus (virtual human) and the responses one would apply to a virtual human and not to a human (Virtual Human Categorization in Figure 1). For example, experience with virtual humans might lead a person to realize how the technology is still far from perfectly mimicking a human. This would form an exclusive virtual human category, where its disposition is non-human and non-social. A negative emotional experience with a virtual human due to its asocial characteristic might expedite this process. For example, one may feel frustrated at an improper emotional expression by a virtual human. If one is an expert on virtual humans, then one would build up a separate set of inference rules that apply to virtual humans exclusively and independently of the rules one applies to humans. For example, a lag in human conversation might mean that the person is simply pondering or choosing their words carefully, but a lag in a virtual human conversation might mean to an expert that the system fell in an infinite loop, or the lag is a result of lack of processing power.

We hypothesize that as long as a clearly observable human-like characteristic of a virtual human exists, the number of human-like characteristics might affect the strength of social response. The majority of psychological studies of categorization have used cues that vary on a few (2–4) dimensions (e.g., shapes, colors; [72]) with most target items ranging from living entities (e.g., cats, dogs) to nonliving entities (e.g., chair, sofa). Given that all prominent categorization models (exemplar, decision bound, and neural network models) assume that the cues may be presented as points in a multidimensional space, the number of cues matters, especially when classifying an entity that has a competing entity (e.g., distinguishing an alligator from a crocodile).

One underlying psychological construct related, believability (i.e., behavioral realism), is relevant here. Although realism seems to strengthen social responses, there are caveats. For example, Bailenson, Yee, Merget and Schroeder [73] investigated behavior and visual appearance realism of avatars and revealed that people disclosed more information to avatars that had low realism. In addition, people expressed their emotions more freely when their avatar did not display emotions. This decrement of social presence is probably due to a mismatch between the reality and expectation of higher realism [74].

All of our human-like characteristics in our framework may contribute to believability (i.e., behavioral realism), and this might increase or decrease the strength of social response which is contingent on the nature of the tasks. To provide clarity in understanding the cognitive mechanics of social responses, we decided to focus on the mere presence of a virtual human as well as the additive effect of human-like characteristics. We found two social phenomena, social facilitation and politeness norm, to be well established in social psychology and suitable to test our hypothesis with a virtual human.

2. Experiment #1: Social Facilitation

Humans behave differently and presumably process information differently, when there is someone else nearby versus when they are alone. More specifically, they tend to do easier tasks better in the presence of others (relative to being alone) and to do difficult tasks less well in the presence of others (relative to being alone). This is referred to as the social facilitation phenomena in social psychology [75] (for a review, see [76]). Prior work demonstrated that virtual humans can produce the social facilitation effect [36]. That is, for easy tasks, performance in the virtual human condition was better than in the alone condition, and for difficult tasks, performance in the virtual human condition was worse than in the alone condition. The social inhibition effect, an opposite of facilitation, was also demonstrated in the presence of virtual humans [77]. For the present study, we were interested in the strength of social facilitation as a function of varying kinds of human-like characteristics. Therefore, participants conducted a series of cognitive tasks alone, in the presence of a human observer, in the presence of a virtual human with varying human-like characteristics, and in the presence of a graphical shape on a computer screen as another control condition.

2.1. Tasks

The experimental tasks needed both breadth and depth to test the social facilitation effect but needed to be applicable to the realm of virtual humans. It seems likely that virtual humans would be most helpful with high level cognitive tasks [78]. Some tasks can be opinion-like (e.g., choosing what to bring on a trip), and others can be more objective (e.g., implementing edits in a document). In addition, the present study examined differences in task performance between simple and complex tasks, so we sought tasks for which it would be possible to produce both easy and difficult instances.

The present study used the following three cognitive tasks: anagrams, mazes, and modular arithmetic. These three tasks provided a good mixture of verbal, spatial, mathematical, and high-level problem solving skills. All three tasks were cognitive tasks and had an objective, and therefore they were within the range of tasks with which a virtual human might assist.

2.1.1. Anagram Task

Social facilitation in anagram tasks has been studied in the context of electronic performance monitoring (EPM), a system whereby every task performed through an electronic device may be analyzed by a remotely located person [79]. The social facilitation effect was clearly observed in the presence of EPM. In the present study, anagrams were divided into two categories (easy or difficult) using normative solution times from Tresselt and Mayzner's [80] anagram research (see also [79]).

2.1.2. Maze Task

Research has suggested that participants tend to perform better in the presence of a human on simple maze tasks [81]. In the present study, simple mazes included wide paths and few blind alleys so that the correct route was readily perceivable, whereas difficult mazes included narrow paths with many blind alleys. Materials for the maze task were similar to the ones of Jackson and Williams [82]. Participants were given a maze and a cursor on the screen and were asked to draw a path to the exit.

2.1.3. Modular Arithmetic Task

The object of Gauss's modular arithmetic is to judge if a problem statement, such as " $50 \equiv 22 \pmod{4}$," is true. In this case, the statement's middle number is subtracted from the first number (i.e., $50 - 22$) and the result of this (i.e., 28) is divided by the last number (i.e., $28 \div 4$). If the quotient is a whole number (as here, 7), then the statement is true.

Beilock et al. [83] claimed that modular arithmetic is advantageous as a laboratory task because it is unusual and, therefore, its learning history can be controlled. In the modular arithmetic tasks, problem statements were given to the participants. Easy problems consisted of single-digit no-borrow subtractions, such as " $7 \equiv 2 \pmod{5}$ "; hard problems consisted of double-digit borrow subtraction operations, such as " $51 \equiv 19 \pmod{4}$ ".

2.2. Participants

One hundred eight participants were recruited from the Georgia Institute of Technology. Participants were compensated with course credit. The study received institutional review board approval.

2.3. Materials

Participants did all tasks (anagrams, maze, and modular arithmetic) on a computer. Java application and Java script were used to implement the tasks on the computer. An additional computer was used to present the virtual human and the graphical shape. For the facial virtual human condition, Haptik Corporation's 3-D character was loaded on this computer (see Figure 2a). The appearance of the virtual human was held constant. The character displayed lifelike behaviors, such as breathing, blinking, and other subtle facial movements. The graphical shape rotated slowly during the experiment (see Figure 2b).

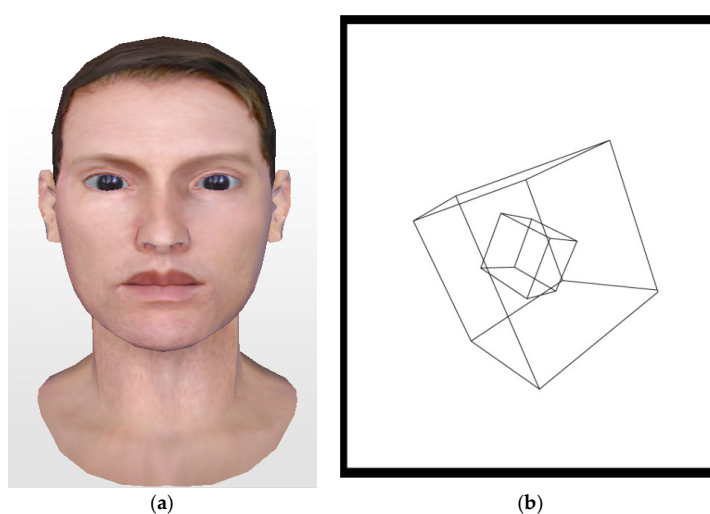


Figure 2. Virtual human (a) and the graphical shape (b) in the present study.

The monitor was positioned so that the virtual human or the graphical shape was oriented to the task screen, not to the participant, and was located about 4 feet (~1.2 m) from the task monitor and about 3.5 feet (~1 m) from the participant. This is also the location where the human observer would sit (see Figure 3).

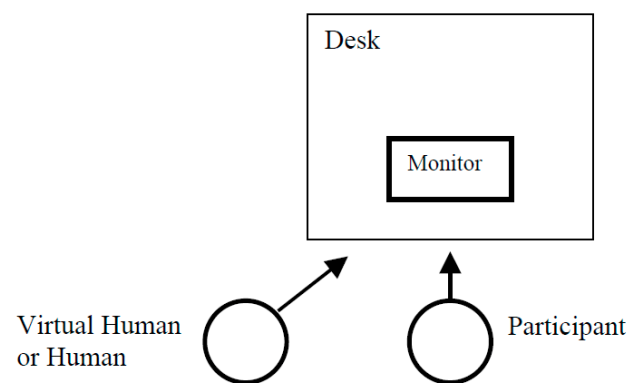


Figure 3. The layout of the experiment.

2.4. Design and Procedure

The present study was a 6 (virtual human type: facial appearance only, voice recordings only, voice synced facial appearance, facial appearance conveying emotion, voice recordings conveying emotion, voice synced facial appearance conveying emotion) \times 3 (task type: anagram task, maze task, modular arithmetic task) \times 2 (task complexity: easy, hard) \times 4 (presence: alone, presence of a human, presence of a virtual human, and presence of a graphical shape) mixed design. We extended the previous study on the social facilitation effects of virtual humans [36] by adding the virtual human variable with six types of manipulation.

All participants did both easy and hard versions of all task types in all presence conditions. Only the virtual human type was a between-subject factor, which varied the virtual human's characteristics. Each participant experienced one of the virtual human types in Table 1. Conditions without conveying emotion conveyed neutral facial expression or neutral voice. Note that, although an underlying model for an emotional expression is necessary for an effective emotional expression, because we are investigating the mere presence effect of emotion, we decided to limit our conditions to conveying an emotion (happy) or non-conveying an emotion (neutral). Voice recordings only condition did not have a facial appearance but only voice implementation.

Table 1. The types of virtual humans manipulated in the social facilitation experiment.

The Type of Virtual Human	Facial	Voice	Emotion	Cover Story
A: Facial appearance only	Yes	No	Neutral	By text
B: Voice recordings only	No	Yes	Neutral	By voice
C: Voice synced facial appearance	Yes	Yes	Neutral	By voice
D: Facial conveying emotion	Yes	No	Happy	By text
E: Voice conveying emotion	No	Yes	Happy	By voice
F: Voice synced facial conveying emotion	Yes	Yes	Happy	By voice

The presence of a graphical shape depicts a minimal visual interaction with the participant to make sure that any effect of a virtual human was due to its human-like characteristics and not due to the mere presence of a graphical component.

The two within subjects factors (task complexity, presence) were crossed to produce eight types of trials, in which participants did:

- a simple task alone
- a simple task in the presence of a human
- a simple task in the presence of a virtual human
- a simple task in the presence of a graphical shape
- a complex task alone

- a complex task in the presence of a human
- a complex task in the presence of a virtual human
- a complex task in the presence of a graphical shape

Every participant experienced multiple instances of each of the eight trial types. The order of the presence factor was varied across participants using a Latin square. That is, some participants did the first set of tasks in the presence of a human (H), the next set in the presence of the graphical shape (G), the third set in the presence of the virtual human (V), and the fourth set alone (A). The four possible orders were:

- $H \rightarrow G \rightarrow V \rightarrow A$
- $G \rightarrow V \rightarrow A \rightarrow H$
- $V \rightarrow A \rightarrow H \rightarrow G$
- $A \rightarrow H \rightarrow G \rightarrow V$

Within a particular presence situation (e.g., virtual human), participants did a block of anagrams, a block of mazes, and a block of modular arithmetic problems. Task order was manipulated using a Latin square resulting in three possible orders:

- anagram \rightarrow maze \rightarrow modular arithmetic
- maze \rightarrow modular arithmetic \rightarrow anagram
- modular arithmetic \rightarrow anagram \rightarrow maze

Within each task block, participants conducted a combination of easy and hard trials for that particular task (e.g., anagrams). The number of easy and hard trials was the same in each block; however, the order of easy and hard trials was one of the three predetermined pseudo-randomized orders.

In the anagram tasks, a five-letter anagram appeared on the screen, and the participants were asked to solve the anagram quickly and accurately by typing in the answer using the keyboard and then pressing the Enter key. Completion time and error rates were measured. Feedback was not given.

In the maze tasks, a maze appeared on the screen. Participants were asked to move the cursor by dragging the mouse through each maze and to find the exit as quickly as possible. Completion time was measured.

In modular arithmetic tasks, a problem statement, such as " $50 \equiv 20 \pmod{4}$," appeared on the screen. Participants were asked to indicate whether the statement was true or false by pressing the corresponding button (Y for "true," N for "false") on the keyboard. Completion time and error rates were measured. Feedback was not given.

Each participant was briefed on each task. Briefing consisted of a demonstration by the experimenter and four hands-on practice trials for the participants so that they could familiarize themselves with the computer and the task.

For conditions involving a human, a virtual human, or a graphical shape, the participants were told that a human, a virtual human, or a graphical shape, was there to "observe" the task, not the participant. Specifically, when a human was present, participants were told, "An observer will be sitting near you to observe the tasks you will doing. The observer will be present to learn more about the tasks and try to catch any mistakes we made in creating the tasks. The observer is not trying to learn how you go about working on the tasks and, in fact, will not be allowed to communicate with you while he is sitting here". When a virtual human was present, the same cover story was delivered by text or by voice depending on the type of virtual human (Table 1).

When a virtual human with only facial appearance (type A in Table 1) or with only facial appearance conveying emotion (type D) or a graphical shape was present, participants were told by the experimenter, "A virtual human (graphical shape) will observe the task. The virtual human (graphical shape) is an artificial intelligence that attempts to analyze events that happen on the computer screen. The virtual human (graphical shape) will be present to learn more about the tasks and try to catch any mistakes we made in creating the tasks. The virtual human (graphical shape) is not trying to learn how you go about

working on the tasks and, in fact, will not be allowed to communicate with you while it is present.”

The virtual human presented the passage with its own voice for the rest of the virtual human types (type B, C, E, and F in Table 1). The CSLU (Center for Spoken Language and Understanding) Toolkit was used to create a computer generated baseline which the human recorder of the voice can reference. The setting of $F_0 = 184$ Hz, pitch range = 55 Hz, and word rate = 198 words per minute for the happy voice (type E and F) and $F_0 = 137$ Hz, pitch range = 30 Hz, and word rate = 177 words per minute for the neutral voice (type B and C) were used. These criteria are from the literature on the markers of emotion in speech [84]. Manipulation checks indicated that the happy recording was recognized as happy voice and the neutral recording was recognized as neutral voice.

Emotion was also conveyed in facial appearances (type D and F; see Figure 4 for comparison in terms of appearance). Although Haptek’s agent software provided a means to manipulate basic emotion (e.g., happy, sad, anger), it was not based on any emotional model or literature, thus manipulation checks were conducted. Participants indicated that the happy facial expression was recognized as happy and the neutral facial expression was recognized as neutral.

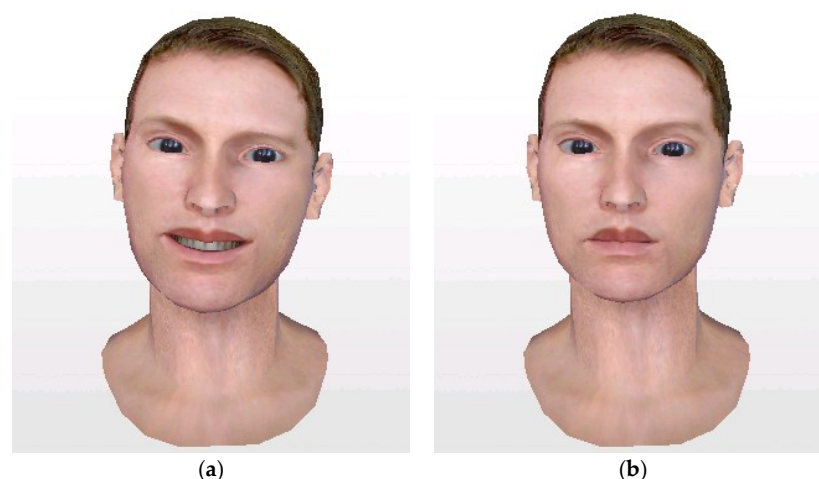


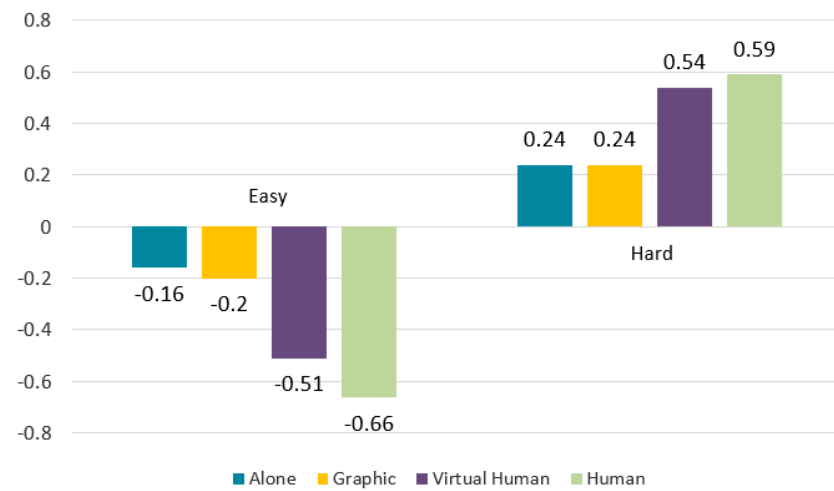
Figure 4. Screenshot of the virtual human with a happy (a) and a neutral (b) expression.

2.5. Results

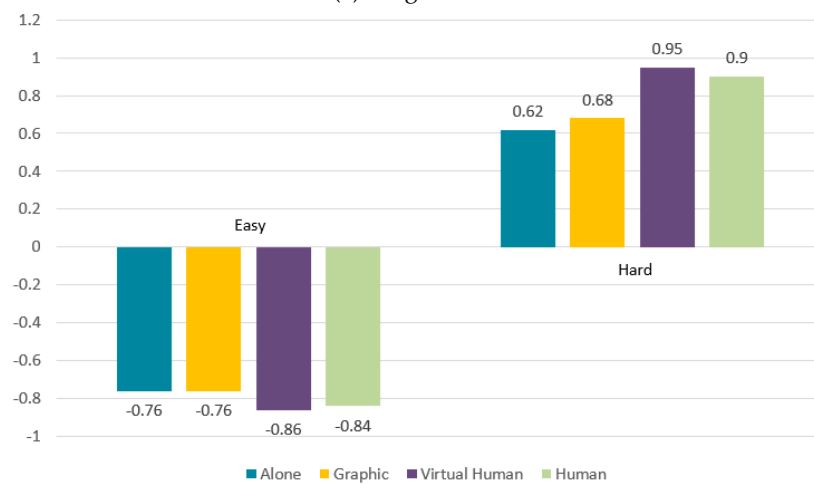
A four-way repeated measures ANOVA (virtual human factor, task factor, complexity factor, and presence factor) was initially conducted and was followed by post hoc analyses. Analysis was conducted on completion time because it was the most frequently measured dependent variable in social facilitation research [85]. The pseudo order and Latin square factors were tested to examine whether they had an effect on performance. Pseudo order and the Latin square orders had no effect on task performance, and all results were collapsed over these variables.

Data were transformed into z scores for each task to perform the analysis involving complexity, presence, and task. The results (summarized in Figure 5) show that the effect of presence on completion time was conditional upon the combination of the task, task complexity, and virtual human type, resulting in a significant four-way interaction of Presence \times Task Type \times Complexity \times Virtual Human Type, $F(30, 612) = 1.54$, $MSE = 0.26$, $p < 0.05$. Of particular importance, the results show that combined across task types (anagram, maze, and modular arithmetic), if the task was easy, completion times in the presence of the virtual human and the real human tended to be faster than in their absence or the presence of the graphical shape, whereas if the task was hard, the mean completion times were slower in the presence of the virtual human and the real human than in their absence or the presence of the graphical shape. This observation is supported by a Presence

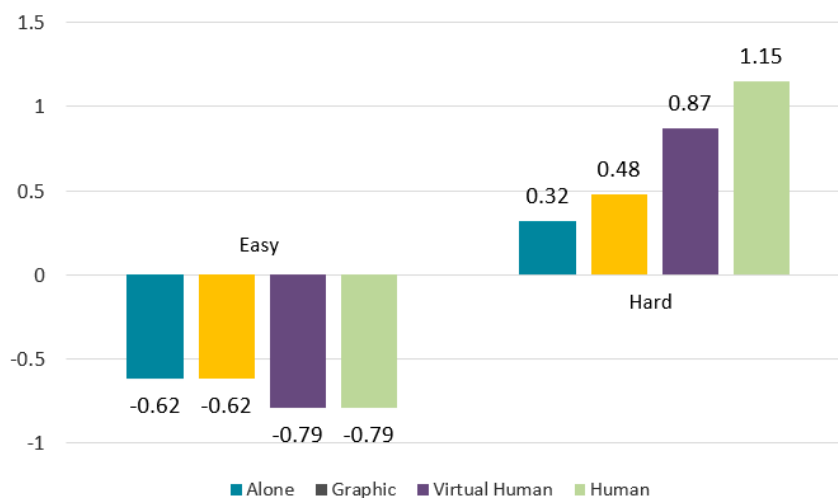
× Complexity interaction that is consistent with the social facilitation effect, $F(3, 306) = 75.65$, $MSE = 0.27$, $p < 0.001$.



(a) anagram task



(b) maze task



(c) modular arithmetic task

Figure 5. Mean completion time in z-score seconds for each task ($n = 108$ each).

There was also a main effect of complexity (easy, hard), $F(1, 102) = 872.36$, $MSE = 1.18$, $p < 0.001$, and presence (alone, virtual human, human, graphical shape), $F(3, 306) = 10.32$, $MSE = 0.25$, $p < 0.001$ but no main effect of task type, $F < 1$.

2.5.1. Post Hoc Analyses for Each Task Type

A three-way Presence \times Task \times Complexity interaction, $F(6, 612) = 3.89$, $MSE = 0.26$, $p < 0.005$, suggests that each task type should be further analyzed for the relationship between presence and complexity. We conducted a post hoc Dunnett's test to compare each presence condition to the alone condition separately for each task type. For each task type, the social facilitation effect for virtual humans was demonstrated and, for each task type, the social facilitation effect for humans was demonstrated. The analyses for each of these observations are presented next.

Virtual human versus alone: For all three tasks (anagram, maze, and modular arithmetic), pairwise comparisons show that completion time for easy tasks was shorter in the virtual human condition than in the alone condition, and that completion time for hard tasks was longer in the virtual human condition than in the alone condition (see Figure 5); anagram easy: $t(214) = 5.07$, $p < 0.001$, $d = -0.37$; anagram hard: $t(214) = 4.86$, $p < 0.001$, $d = 0.28$; maze easy: $t(214) = 2.00$, $p < 0.05$, $d = -0.05$; maze hard: $t(214) = 4.14$, $p < 0.001$, $d = 0.44$; modular arithmetic easy: $t(214) = 2.43$, $p < 0.05$, $d = -0.47$; modular arithmetic hard: $t(214) = 7.86$, $p < 0.001$, $d = 0.84$.

Human versus alone: For all three tasks, pairwise comparisons show that completion time for easy tasks was shorter in the human condition than in the alone condition, and completion time for hard tasks was longer in the human condition than in the alone condition (see Figure 5); anagram easy: $t(214) = 7.00$, $p < 0.001$, $d = -0.49$; anagram hard: $t(214) = 4.86$, $p < 0.001$, $d = 0.33$; maze easy: $t(214) = 2.23$, $p < 0.05$, $d = -0.33$; maze hard: $t(214) = 4.68$, $p < 0.001$, $d = 0.53$; modular arithmetic easy: $t(214) = 2.42$, $p < 0.05$, $d = -0.47$; modular arithmetic hard: $t(214) = 11.72$, $p < 0.001$, $d = 1.26$.

2.5.2. Post Hoc Analyses for Each Virtual Human Type

Voice synced facial appearance versus human: Of interest was whether the social response to a virtual human with voice synced facial appearance is similar to social response to a human. For all three tasks, pairwise comparisons show that completion time for both easy and hard tasks did not significantly differ from the human condition *with the exception of the maze hard task*; anagram easy: $t(34) < 1$; anagram hard: $t(34) < 1$; maze easy: $t(34) < 1$; maze hard: $t(34) = 4.17$, $p < 0.01$, $d = 0.57$; modular arithmetic easy: $t(34) < 1$; modular arithmetic hard: $t(34) < 1$ (see Table 2). We acknowledge that these are primarily null results, therefore we conducted a post hoc power analysis with the program G*Power [86] with power set at 0.8 and $\alpha = 0.05$, $d = 0.5$, two-tailed. The results suggest an n of approximately 384 would be needed to achieve appropriate statistical power; the limitation of these findings will be discussed.

Table 2. Post hoc analyses for each virtual human type.

The Type of Virtual Human	Anagram Easy	Anagram Hard	Maze Easy	Maze Hard	Modular Arithmetic Easy	Modular Arithmetic Hard
Facial appearance only	$p < 0.05$	$p < 0.01$	$p < 0.05$	$p < 0.05$	$t < 1$	$t < 1$
Voice recordings only	$p < 0.05$	$p < 0.05$	$p > 0.05$ ($t = 1.17$)	$p < 0.05$	$p < 0.05$	$p < 0.01$
Voice synced facial appearance	$t < 1$	$t < 1$	$t < 1$	$p < 0.01$	$t < 1$	$t < 1$
Facial appearance conveying emotion	$t < 1$	$p < 0.05$	$p < 0.05$	$p < 0.05$	$t < 1$	$p < 0.05$
Voice recordings conveying emotion	$t < 1$	$p < 0.05$	$p < 0.05$	$p < 0.05$	$t < 1$	$p < 0.05$
Voice synced facial appearance conveying emotion	$t < 1$	$t < 1$	$t < 1$	$t < 1$	$t < 1$	$t < 1$

Facial appearance versus human: Of interest was whether the social response to a face only virtual human is similar to the social response to a human. For all three tasks, pairwise comparisons show that completion time for both easy and hard tasks was statistically different from the human condition *with the exception of the modular easy task and the hard task*;

anagram easy: $t(34) = 2.5, p < 0.05, d = 0.1$; anagram hard: $t(34) = 3.0, p < 0.01, d = -0.26$; maze easy: $t(34) = 2.47, p < 0.05, d = 2.23$; maze hard: $t(34) = 2.0, p < 0.05, d = -0.19$; modular arithmetic easy: $t(34) < 1$; modular arithmetic hard: $t(34) < 1$.

Voice recordings versus human: Of interest was whether the social response to a voice only virtual human is similar to the social response to a human. For all three tasks, comparisons show that completion time for both easy and hard tasks was statistically different from the human condition *with the exception of the maze easy task*; anagram easy: $t(34) = 3.25, p < 0.05, d = 0.41$; anagram hard: $t(34) = 2.25, p < 0.05, d = -0.56$; maze easy: $t(34) = 1.17, p > 0.05, d = 0.65$; maze hard: $t(34) = 2.50, p < 0.05, d = -0.17$; modular arithmetic easy: $t(34) = 2.13, p < 0.05, d = -0.003$; modular arithmetic hard: $t(34) = 3.00, p < 0.01, d = -0.66$.

Facial appearance conveying emotion versus human: Of interest was whether the social response to a facial appearance conveying emotion is similar to the social response to a human. For all three tasks, comparisons show that completion time for both easy and hard tasks was statistically different from the human condition *with the exception of the anagram easy and the modular arithmetic easy*; anagram easy: $t(34) < 1$; anagram hard: $t(34) = 2.25, p < 0.05, d = 0.07$; maze easy: $t(34) = 2.41, p < 0.05, d = 0.72$; maze hard: $t(34) = 2.17, p < 0.05, d = -0.22$; modular arithmetic easy: $t(34) < 1$; modular arithmetic hard: $t(34) = 2.10, p < 0.05, d = -0.51$.

Voice recordings conveying emotion versus human: Of interest was whether the social response to a voice recording conveying emotion is similar to the social response to a human. For all three tasks, pairwise comparisons show that completion time for both easy and hard tasks was statistically different from the human condition *with the exception of anagram easy and modular arithmetic easy*; anagram easy: $t(34) < 1$; anagram hard: $t(34) = 2.17, p < 0.05, d = 0.02$; maze easy: $t(34) = 2.23, p < 0.05, d = 0.45$; maze hard: $t(34) = 2.41, p < 0.05, d = -0.05$; modular arithmetic easy: $t(34) < 1$; modular arithmetic hard: $t(34) = 2.10, p < 0.05, d = -0.01$.

Voice synced facial appearance conveying emotion versus human: Of interest was whether the social response to a voice synced facial appearance conveying emotion is similar to the social response to a human. For all three tasks, pairwise comparisons show that completion time for both easy and hard tasks was not significantly different from the human condition; anagram easy: $t(34) < 1$; anagram hard: $t(34) < 1$; maze easy: $t(34) < 1$; maze hard: $t(34) < 1$; modular arithmetic easy: $t(34) < 1$; modular arithmetic hard: $t(34) < 1$.

In general, the social facilitation experiment showed that as with a human, virtual humans do produce social facilitation. In addition, performance in the presence of voice synced facial appearance and voice synced facial appearance conveying emotion seems to elicit stronger social facilitation (i.e., no statistical difference compared to performance in the human presence condition) than in the presence of voice only or face only.

3. Experiment #2: Politeness Norm

Thus far, data from the social facilitation experiment showed that as with a human, virtual humans can produce social facilitation with all virtual human type conditions. In addition, combining voice with face seems to elicit stronger social facilitation (i.e., no statistical difference compared to performance in the human presence condition) than in the presence of voice only or face only. Conveying emotion did not produce this additional effect. The purpose of the second experiment was first, to expand this result to a different social behavior known in the literature as the “politeness norm”.

In the social psychology literature, socially desirable response bias in interview situations has been well documented [87–89]. One type of social desirability effect occurs when people bias their responses according to the perceived preference of the interviewer. For example, when person A is asked by person B to evaluate person B (direct evaluation), person A generally applies the politeness norm and gives a positive response to avoid offending Person B. On the other hand, if person C were to ask person A to evaluate person

B (indirect evaluation), then person A tends to give a more honest (likely more negative) response to Person C.

Studies by Nass [14,16] on whether people apply the politeness norm to computers provided some insights in designing the present study. In the present experiment, of interest is whether direct, as opposed to indirect, requests for evaluations of virtual humans elicit more positive responses. In addition, this experiment collected data of people's responses to a human presenter as a basis for comparison.

Participants were presented and quizzed either by a human presenter or a virtual human presenter that varied in terms of features. Participants then evaluated the presenter's performance either by responding directly to the presenter or indirectly via a paper and pencil questionnaire.

3.1. Participants

Eighty participants were recruited from the Georgia Institute of Technology. Participants were compensated with course credit. The study received institutional review board approval.

3.2. Materials

Participants did the experiment on a computer. HTML was used to implement the tasks on the computer. Audio (voice only condition) or video clips (facial appearance only and face with voice conditions) of Haptik Corporation's 3-D character were imbedded in the HTML code.

3.3. Design and Procedure

The present study was a 4 (presenter: virtual human with facial appearance only, virtual human with voice only, virtual human with voice synced facial appearance, human) \times 2 (evaluation: direct, indirect) between-subject design. The virtual human conditions were identical to the social facilitation experiment except conditions that conveyed emotion (see Table 3). Such conditions were eliminated because conveying emotion did not produce additive effects in the first experiment.

Table 3. The types of virtual humans manipulated in the politeness norm experiment.

The Type of Virtual Human	Facial Appearance	Voice Implementation
A: Facial appearance only	Yes	No
B: Voice recordings only	No	Yes
C: Voice synced facial appearance	Yes	Yes

The experiment consisted of four sessions: presentation session, testing session, scoring session, and interview session. When they arrived, participants were briefed on the general procedure of the experiment with a prepared script.

3.3.1. Presentation Session

In this session participants were told that they will learn facts about national geography. The presenter was either a virtual human with varying characteristics (see Table 3) or a human presenter based on the condition each participant was assigned to. The presenter presented each participant with 20 facts (e.g., The tallest waterfall is Angel Falls in Venezuela; see Appendix A). The participant then answered a question about whether they knew about the fact by clicking a button labeled with "I knew the fact" or a button labeled with "I didn't know the fact". Participants were told that they would receive 20 out of 1000 facts and that facts will be chosen based on their familiarity of a subject as assessed by their performance on earlier facts. This was to provide a sense of interactivity with the virtual humans [90]. In reality, all participants received the same 20 facts in the same order.

3.3.2. Testing Session

The presenter then administered a 12-item multiple choice test with 5 options per item. The participants were told that 12 questions (e.g., The tallest waterfall is in ____; see Appendix B) would be randomly drawn from a total of 5000 questions. In reality, all participants received the same 12 questions in the same order.

3.3.3. Scoring Session

During this session, the presenter provided the participant's score as well as an evaluation on its own performance. All participants were informed that they had answered the same 10 of 12 questions correctly and received identical and positive evaluation of the presenter ("The presenter performed well"). A constant and fixed score (10 out of 12) was presented to prevent individual scores from affecting the evaluation and become a confounding variable. However, it was also important to not let participants realize the scoring was fake. Facts and questions were designed so that the score was acceptable by the participants. Manipulation checks indicated participants were not surprised with the score and believed it.

3.3.4. Interview Session

Following the completion of the scoring session, participants were asked 13 questions about the performance of the presenter (e.g., competence, likable, fair). If the presenter was a virtual human, then this interview was conducted by either the same virtual human (direct evaluation) or through a paper-and-pencil questionnaire (indirect evaluation). If the presenter was a human, then the interview was conducted by the same human presenter (direct evaluation) or through a paper-and-pencil questionnaire (indirect evaluation).

3.4. Results

We averaged all 13 items in the interview questionnaire to a single score for each participant to capture each participant's overall valence toward the presenter. Consistent with the interview questionnaire, higher scores indicated more positive responses toward the presenter.

A two-way between subjects ANOVA (presenter and evaluation factor) was conducted and was followed by a post hoc Dunnett's test. Analysis was conducted on the overall valence toward the presenter (i.e., the average of all 13 items). The results (summarized in Figure 6) show that combined across presenter types, consistent with the politeness prediction, participants gave more positive responses when the presenter asked about its own performance (direct evaluation) as compared to when participants answered on paper-and-pencil (indirect evaluation). This observation was supported by a main effect of Evaluation (direct, evaluation), $F(1, 79) = 10.27$, $MSE = 12.74$, $p < 0.01$. There was also a main effect of Presenter (virtual human with facial appearance only, virtual human with voice only, virtual human with voice synced facial appearance, human), $F(3, 237) = 11.94$, $MSE = 14.81$, $p < 0.001$, but no Evaluation by Presenter interaction, $F < 1$.

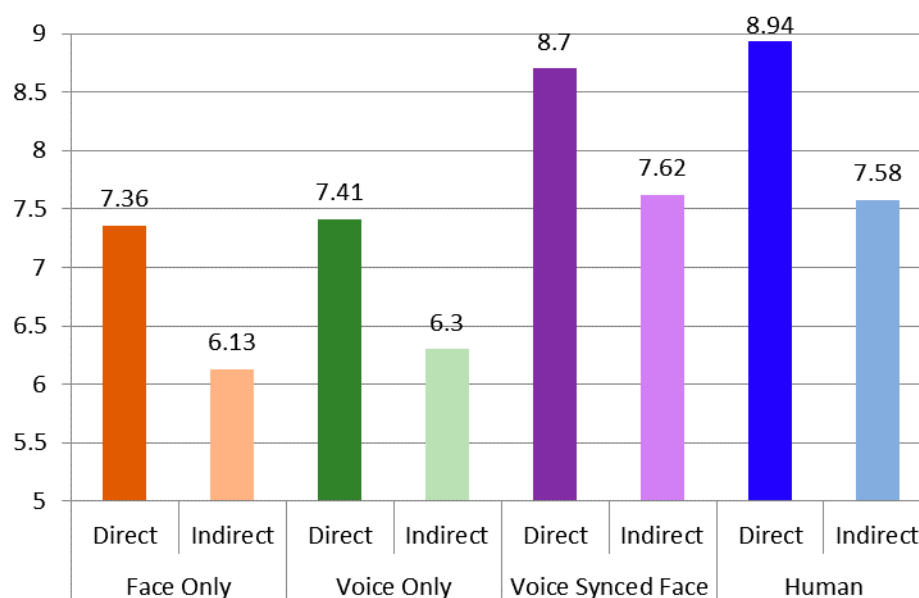


Figure 6. Mean valence toward the presenter for each evaluation type. (direct, indirect) for each presenter type (face only, voice only, voice synced face, human) ($n = 80$).

Post Hoc Analyses

Direct versus indirect: The main effects of Presenter and Evaluation suggest that each presenter type should be further analyzed to see whether participants applied the politeness norm to all presenter types. We conducted a post hoc Dunnett's test to compare each direct condition to the indirect condition separately for each presenter type. For each presenter type, the politeness norm was demonstrated; face only: $t(72) = 2.46$, $p < 0.05$, $d = 0.57$; voice only: $t(72) = 2.22$, $p < 0.05$, $d = 0.52$; voice synced face: $t(72) = 2.16$, $p < 0.05$, $d = 0.54$; human: $t(72) = 2.72$, $p < 0.01$, $d = 0.73$.

Virtual human versus human: Of interest was whether the social response to the virtual human was similar to that of the social response to the human. The valence toward the voice synced face condition did not differ significantly from the valence toward the human presenter condition for both direct and indirect conditions, $t < 1$. However, this valence toward the voice synced face voice was significantly higher than the valence toward the voice only condition for both direct and indirect conditions, $t(72) = 2.6$, $p < 0.05$. There was no significant difference between the voice-only condition and the face only condition for both direct and indirect conditions, $t < 1$.

As with the previous experiment, we conducted a post hoc power analysis with the program G*Power with power set at 0.8 and $\alpha = 0.05$, $d = 0.5$, two-tailed. The results suggest an n of approximately 640 would be needed to achieve appropriate statistical power; once again the limitation of these findings will be discussed.

Results from the politeness experiment indicate that, consistent with the politeness norm, participants who evaluated their presenter directly reported the presenter's performance more favorably than participants who evaluated their presenter indirectly. This was observed not only with the human presenter but with virtual human presenters in all four conditions. In addition, consistent with the social facilitation experiment, the valence toward the voice synced facial appearance was not significantly different from the valence toward the human presenter condition. In other words, when voice is combined with face (i.e., voice synced face) an additive effect seems to occur where the social responses to virtual humans with such features are indistinguishable from the social responses to humans.

4. Discussion and Conclusions

The purpose of this research was to gain a deeper understanding of the social dimension of the interaction between users and virtual humans. We propose that human-like characteristics of virtual humans act as a cue that leads a person to place the agent into the category “human” and thus, elicits social responses.

Given this framework, experiments were designed to answer whether a virtual human elicits stronger social responses as that virtual human is provided with more human-like characteristics. The first experiment investigated whether virtual humans can evoke social facilitation. The second experiment investigated whether people apply politeness norms to virtual humans. Results from both studies demonstrated that virtual humans can elicit social behaviors. In addition, performance in the presence of voice synced facial appearance seems to elicit stronger social responses than in the presence of voice only or face only. However, we acknowledge that a larger N would have been needed to achieve appropriate power to completely rule out the false positive.

Overall the results suggest that there are minimal requirements for humans to classify a non-biological entity as a human and respond socially (See Figure 7). Voice and facial expression appear to be such characteristics. In addition, performance in the presence of voice synced facial appearance seems to elicit stronger social responses than in the presence of voice only or face only. The pattern of results suggests that conveying emotion did elicit the social facilitation effect. The results seem to support Nass’ early speculation that the more an agent presents human-like characteristics, the more likely they will elicit social behavior [18].

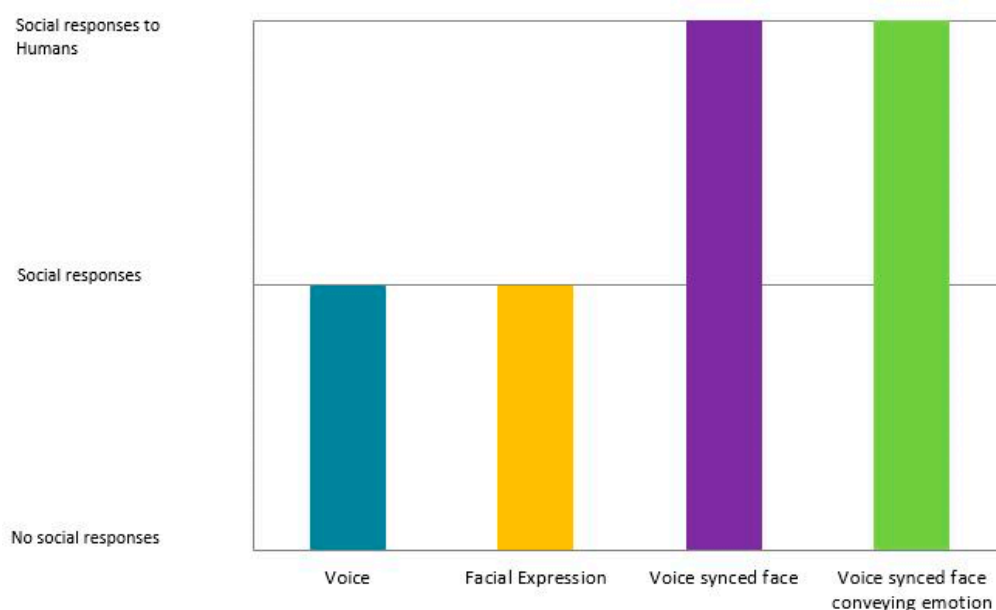


Figure 7. Social responses as a function of human-like characteristics.

The most recent review of research on social presence generally supports our findings [91]. Studies found that participants perceived the lowest level of social presence when communicating via text-based medium compared to richer media [92–94]. Nevertheless, such non-human mediated interaction had elicited a lower level of social presence compared to the real human presence [91]. The current study had limited dependent measures (e.g., completion time), but future studies should measure social presence to assess the relationship with task performance. History and length of interaction are critical factors in any social interaction as depicted in our framework, so further research on types of social norms and behaviors and the long-term effects of social interactions with virtual humans is required.

As mentioned earlier, the inherent characteristics of the experiment design limit the generalization of the study. Because our study focused on the “mere presence” effect of the virtual human, emotion, as a variable, was explicitly implemented as task-irrelevant displays of emotions. Nevertheless, the importance of emotional expression related to social effects is well-established in the literature [95,96]. Specifically, emotions are known to communicate task-relevant information [97–99]. To achieve ecological validity, future research requires manipulation of emotional expression that is relevant to the task and includes participants’ moment-to-moment reactions.

This caveat may be further exacerbated in a cross-cultural domain involving both the expresser (virtual human) and the perceiver (human). We now have some understanding as to how emotional expression and perception have different social effects on different cultural members. For example, western negotiators perceive anger expression as appropriate and therefore provide larger concessions than eastern negotiators [100]. Adam and Shirako [101] further found that this display of anger elicited greater cooperative efforts when the expressers were eastern negotiators because they are perceived as more threatening compared with angry western negotiators. This effect is mediated by the counterparts’ stereotype of East Asians being emotionally inexpressive [102]. Easterners project onto others the emotion that the generalized other would feel than while Westerners are more likely to project their own specific emotions [103], further proof that the former group is more interdependent than individualistic relative to the latter group [104]. There is also empirical evidence that ethnic stereotyping holds even with computers; that is, people perceive a same-ethnicity virtual human to be more attractive, trustworthy, persuasive, and intelligent [35].

Such findings have a profound impact on the design of virtual humans. It may not be wise to deploy globally a virtual human with the same appearance and interaction, because the perceiver of emotional expression is different by region and the intended social effect would entail different consequences. Distinct virtual humans for different regions are further justified because the tasks and execution of them would be loaded with cultural meaning that varies by region.

There are other potential human-like factors (e.g., signs of intelligence, language) that future work might consider. In addition, future work can explore more deeply the claim that once the “human” categorization is made, asocial cues are ignored. While the present results are consistent with this claim, they do not directly test it.

The study is limited to virtual humans that may have a task-based role. However, a significant number of agents are designed to develop trust and rapport for social-emotional relationships with their users. Future work should test the effect with empathic or relational agents. Such agents elicit change in attitudes and behaviors by remembering interactions and managing future expectations [105]. The system should carefully respond to the user’s expectations for the virtual human’s social role. The current findings may turn out differentially with interesting implications.

The present results have implications for the design of information systems involving human-computer interaction. Given the frequent high cost and limited time to design a virtual human, it is important to understand the necessary and minimal requirements to elicit a social response from the user. Results suggest that designers of such systems should consider that users behave differently in the presence of virtual humans, as compared with when they are alone, and that the nature of the behavior depends on the task (such as task difficulty). A design decision to present a virtual human should be a deliberate and thoughtful one. One should use caution when generalizing the present findings. For example, in case of tutoring systems, because the virtual human has the role of collaborator or facilitator of learning, its presence may entail different social effects. Further investigation to effects of roles between observer and collaborator is required.

The results provide insights for commercial virtual agents as well. The success of a virtual human assisting a user to conduct financial transactions depends on how much the user trusts the virtual human; trust is a social behavior. The present study also provides

insights into some practical questions concerning implementation, especially when a designer intends to build a social component into a virtual human and expects social responses from users. For example, given limited resources (e.g., time, money), which characteristic should designers implement first? A clear and identifiable representation of a voice or a face might be enough to elicit a strong and reliable social behavior. A further combination of voice and face implementation might be required for a more complex social reaction from a user. Based on the present study, implementation of emotion might not be necessary to elicit social behaviors. The present study provides a framework for continued systematic investigation of such issues.

Lastly, we would like to conclude with a note on the ethical consequences of virtual humans. Our intention is not to obscure the difference between humans and virtual humans, nor do the findings justify such a claim. Virtual humans or any artificial intelligence should not replace human relationships but rather be used as a tool to promote human social interaction. The social nature of virtual humans does not condone deceit, tricking users into thinking they are interacting with a person [106]. That is, the virtual human's capabilities and limitations should be transparent to the user.

Author Contributions: S.P.: conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing—original draft preparation, R.C.: conceptualization, methodology, writing—review and editing, supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board of the Georgia Institute of Technology (protocol code H08164, approved at 30 June 2008).

Informed Consent Statement: Informed consent was obtained from all participants involved in the study. Written informed consent has been obtained from the participants to publish this paper.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A 20 Facts on World Geography

1. Asia is a larger continent than Africa.
2. The Amazon rainforest produces more than 20% of the world's oxygen supply.
3. The largest desert in Africa is the Sahara desert.
4. There are more than roughly 200 million people living in South America.
5. Australia is the only country that is also a continent.
6. The tallest waterfall is Angel Falls in Venezuela.
7. More than 25% of the world's forests are in Siberia.
8. The largest country in South America is Brazil. Brazil is almost half the size of South America.
9. Europe produces just over 18 percent of all the oil in the world.
10. Asia has the largest population with over 3 billion people.
11. Europe is the second smallest continent with roughly 4 million square miles.
12. The four largest nations are Russia, China, USA, and Canada.
13. The largest country in Asia by population is China with more than 1 billion people.
14. The Atlantic Ocean is saltier than the Pacific Ocean.
15. The longest river in North America is roughly 3000 feet and it is the Mississippi River in the United States.
16. Poland is located in central Europe and borders with Germany, Czech Republic, Russia, Belarus, Ukraine and Lithuania.
17. Australia has more than 28 times the land area of New Zealand, but its coastline is not even twice as long.
18. The coldest place in the Earth's lower atmosphere is usually not over the North or South Poles.
19. The largest country in Asia by area is Russia.

20. Brazil is so large that it shares a border with all South American countries except Chile and Ecuador.

Appendix B 12 Questions on World Geography

1. The nation that has the largest forests is ____.
(1) Brazil (2) Russia (3) Germany (4) USA
2. Which country **does not** border Poland?
(1) Germany (2) Ukraine (3) Slovakia (4) France
3. China takes almost ____% of Asia in terms of population.
(1) 15% (2) 33% (3) 42% (4) 52%
4. What is the third largest nation?
(1) China (2) USA (3) Canada (4) Russia
5. The coldest place in the Earth's lower atmosphere is ____.
(1) North Poles (2) South Poles (3) Equator (4) Siberia
6. Which continent is the largest in the world?
(1) Asia (2) Africa (3) Europe (4) Australia
7. Brazil does share a border with ____?
(1) Chile (2) Colombia (3) Mexico (4) Ecuador
8. The smallest continent is roughly ____ million square miles.
(1) 3 (2) 4 (3) 8 (4) 25
9. The tallest waterfall is in ____.
(1) Europe (2) Asia (3) South America (4) North America
10. Which continent is the smallest in the world?
(1) Europe (2) Australia (3) Africa (4) Asia
11. Which country **does not** have a coastline?
(1) New Zealand (2) Australia (3) Switzerland (4) Austria
12. Brazil is roughly 3,300,000 square miles. Given this, South America is roughly ____ square miles.
(1) 1,000,000 (2) 4,000,000 (3) 6,000,000 (4) 9,000,000

References

1. Jeong, S.; Hashimoto, N.; Makoto, S. A novel interaction system with force feedback between real-and virtual human: An entertainment system: Virtual catch ball. In Proceedings of the 2004 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology, Singapore, 3–5 June 2004; pp. 61–66.
2. Zibrek, K.; Kokkinara, E.; McDonnell, R. The effect of realistic appearance of virtual characters in immersive environments-does the character's personality play a role? *IEEE Trans. Vis. Comput. Graph.* **2018**, *24*, 1681–1690. [[CrossRef](#)]
3. Zorriassatine, F.; Wykes, C.; Parkin, R.; Gindy, N. A survey of virtual prototyping techniques for mechanical product development. *Proc. Inst. Mech. Eng. Part B J. Eng. Manuf.* **2003**, *217*, 513–530. [[CrossRef](#)]
4. Dignum, V. Social agents: Bridging simulation and engineering. *Commun. ACM* **2017**, *60*, 32–34. [[CrossRef](#)]
5. Glanz, K.; Rizzo, A.S.; Graap, K. Virtual reality for psychotherapy: Current reality and future possibilities. *Psychother. Theory Res. Pract. Train.* **2003**, *40*, 55. [[CrossRef](#)]
6. Javaid, M.; Haleem, A. Virtual reality applications toward medical field. *Clin. Epidemiol. Glob. Health* **2020**, *8*, 600–605. [[CrossRef](#)]
7. Hill, R.W., Jr.; Gratch, J.; Marsella, S.; Rickel, J.; Swartout, W.R.; Traum, D.R. Virtual Humans in the Mission Rehearsal Exercise System. *Künstliche Intell.* **2003**, *17*, 5.

8. Roessingh, J.J.; Toubman, A.; van Oijen, J.; Poppinga, G.; Hou, M.; Luotsinen, L. Machine learning techniques for autonomous agents in military simulations—Multum in Parvo. In Proceedings of the 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Banff, AB, Canada, 5–8 October 2017; pp. 3445–3450.
9. Baldwin, K. Virtual avatars: Trans experiences of ideal selves through gaming. *Mark. Glob. Dev. Rev.* **2019**, *3*, 4. [\[CrossRef\]](#)
10. Parmar, D.; Olafsson, S.; Utami, D.; Bickmore, T. Looking the part: The effect of attire and setting on perceptions of a virtual health counselor. In Proceedings of the 18th International Conference on Intelligent Virtual Agents, Sydney, NSW, Australia, 5–8 November 2018; pp. 301–306.
11. Kamińska, D.; Sapiński, T.; Wiak, S.; Tikk, T.; Haamer, R.E.; Avots, E.; Helmi, A.; Ozcinar, C.; Anbarjafari, G. Virtual reality and its applications in education: Survey. *Information* **2019**, *10*, 318. [\[CrossRef\]](#)
12. Achenbach, J.; Waltemate, T.; Latoschik, M.E.; Botsch, M. Fast generation of realistic virtual humans. In Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology, Gothenburg, Sweden, 8–10 November 2017; pp. 1–10.
13. Lester, J.C.; Converse, S.A.; Kahler, S.E.; Barlow, S.T.; Stone, B.A.; Bhogal, R.S. The persona effect: Affective impact of animated pedagogical agents. In Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems, Atlanta, GA, USA, 22–27 March 1997; pp. 359–366.
14. Nass, C.; Moon, Y.; Carney, P. Are people polite to computers? Responses to computer-based interviewing systems 1. *J. Appl. Soc. Psychol.* **1999**, *29*, 1093–1109. [\[CrossRef\]](#)
15. Reeves, B.; Nass, C. *The Media Equation: How People Treat Computers, Television, and New Media Like Real People*; Cambridge University Press: Cambridge, UK, 1996.
16. Nass, C.; Steuer, J.; Tauber, E.R. Computers are social actors. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Boston, MA, USA, 24–28 April 1994; pp. 72–78.
17. Myers, D.G.; Smith, S.M. *Exploring Social Psychology*; McGraw-Hill: New York, NY, USA, 2012.
18. Nass, C.; Moon, Y. Machines and mindlessness: Social responses to computers. *J. Soc. Issues* **2000**, *56*, 81–103. [\[CrossRef\]](#)
19. Turing, A.M.; Haugeland, J. *Computing Machinery and Intelligence*; MIT Press: Cambridge, MA, USA, 1950.
20. Gilbert, D.T.; Fiske, S.T.; Lindzey, G. *The Handbook of Social Psychology*; Oxford University Press: Oxford, UK, 1998.
21. Allport, G.W.; Clark, K.; Pettigrew, T. *The Nature of Prejudice*; Addison-Wesley: Boston, MA, USA, 1954.
22. Fiske, S.T. *Stereotyping, Prejudice and Discrimination*; McGraw-Hill: New York, NY, USA, 1998.
23. Shiffrin, R.M.; Schneider, W. Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychol. Rev.* **1977**, *84*, 127. [\[CrossRef\]](#)
24. Schneider, W.; Shiffrin, R.M. Controlled and automatic human information processing: I. Detection, search, and attention. *Psychol. Rev.* **1977**, *84*, 1–66. [\[CrossRef\]](#)
25. Bargh, J.A. Conditional automaticity: Varieties of automatic influence in social perception and cognition. *Unintended Thought* **1989**, 3–51.
26. Kiesler, S.; Sproull, L. Social human-computer interaction. In *Human Values and the Design of Computer Technology*; Cambridge University Press: Cambridge, MA, USA, 1997; pp. 191–199.
27. Blascovich, J.; Loomis, J.; Beall, A.C.; Swinth, K.R.; Hoyt, C.L.; Bailenson, J.N. Immersive virtual environment technology as a methodological tool for social psychology. *Psychol. Inq.* **2002**, *13*, 103–124. [\[CrossRef\]](#)
28. Von der Pütten, A.M.; Krämer, N.C.; Gratch, J.; Kang, S.H. “It doesn’t matter what you are!” explaining social effects of agents and avatars. *Comput. Human Behav.* **2010**, *26*, 1641–1650. [\[CrossRef\]](#)
29. Hegel, F.; Krach, S.; Kircher, T.; Wrede, B.; Sagerer, G. Understanding social robots: A user study on anthropomorphism. In Proceedings of the RO-MAN 2008-The 17th IEEE International Symposium on Robot and Human Interactive Communication, Munich, Germany, 1–3 August 2008; pp. 574–579.
30. Fox, J.; Ahn, S.J.; Janssen, J.H.; Yeykelis, L.; Segovia, K.Y.; Bailenson, J.N. Avatars versus agents: A meta-analysis quantifying the effect of agency on social influence. *Hum. Comput. Interact.* **2015**, *30*, 401–432. [\[CrossRef\]](#)
31. Gray, H.M.; Gray, K.; Wegner, D.M. Dimensions of mind perception. *Science* **2007**, *315*, 619. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Yee, N.; Bailenson, J.N.; Rickertsen, K. A meta-analysis of the impact of the inclusion and realism of human-like faces on user experiences in interfaces. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, 28 April–3 May 2007; pp. 1–10.
33. Selvarajah, K.; Richards, D. The use of emotions to create believable agents in a virtual environment. In Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems, Utrecht, The Netherlands, 25–29 July 2005; pp. 13–20.
34. Zambaka, C.; Goolkasian, P.; Hodges, L. Can a virtual cat persuade you? The role of gender and realism in speaker persuasiveness. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 22–27 April 2006; pp. 1153–1162.
35. Nass, C.; Isbister, K.; Lee, E.J. Truth is beauty: Researching embodied conversational agents. *Embodied Conversat. Agents* **2000**, 374–402.
36. Park, S.; Catrambone, R. Social facilitation effects of virtual humans. *Hum. Factors* **2007**, *49*, 1054–1060. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Rickenberg, R.; Reeves, B. The effects of animated characters on anxiety, task performance, and evaluations of user interfaces. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, The Hague, The Netherlands, 1–6 April 2000; pp. 49–56.

38. Sproull, L.; Subramani, M.; Kiesler, S.; Walker, J.H.; Waters, K. When the interface is a face. *Hum. Comput. Interact.* **1996**, *11*, 97–124. [\[CrossRef\]](#)
39. Bond, E.K. Perception of form by the human infant. *Psychol. Bull.* **1972**, *77*, 225. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Morton, J.; Johnson, M.H. CONSPEC and CONLERN: A two-process theory of infant face recognition. *Psychol. Rev.* **1991**, *98*, 164. [\[CrossRef\]](#)
41. Diener, E.; Fraser, S.C.; Beaman, A.L.; Kelem, R.T. Effects of deindividuation variables on stealing among Halloween trick-or-treaters. *J. Personal. Soc. Psychol.* **1976**, *33*, 178. [\[CrossRef\]](#)
42. Stern, S.E.; Mullennix, J.W.; Dyson, C.L.; Wilson, S.J. The persuasiveness of synthetic speech versus human speech. *Hum. Factors* **1999**, *41*, 588–595. [\[CrossRef\]](#)
43. Mullennix, J.W.; Stern, S.E.; Wilson, S.J.; Dyson, C.L. Social perception of male and female computer synthesized speech. *Comput. Human Behav.* **2003**, *19*, 407–424. [\[CrossRef\]](#)
44. Lee, K.M.; Nass, C. Social-psychological origins of feelings of presence: Creating social presence with machine-generated voices. *Media Psychol.* **2005**, *7*, 31–45. [\[CrossRef\]](#)
45. Nass, C.; Foehr, U.; Brave, S.; Somoza, M. The effects of emotion of voice in synthesized and recorded speech. In *Proceedings of the AAAI Symposium Emotional and Intelligent II: The Tangled Knot of Social Cognition 2001*; AAAI: North Falmouth, MA, USA, 2001.
46. Nass, C.; Moon, Y.; Green, N. Are machines gender neutral? Gender-stereotypic responses to computers with voices. *J. Appl. Soc. Psychol.* **1997**, *27*, 864–876. [\[CrossRef\]](#)
47. Lee, K.M.; Nass, C. Designing social presence of social actors in human computer interaction. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, Ft. Lauderdale, FL, USA, 5–10 April 2003; pp. 289–296.
48. Eyssel, F.; De Ruiter, L.; Kuchenbrandt, D.; Bobinger, S.; Hegel, F. If you sound like me, you must be more human': On the interplay of robot and user features on human-robot acceptance and anthropomorphism. In *Proceedings of the 2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, Boston, MA, USA, 5–8 March 2012; pp. 125–126.
49. Gratch, J.; Marsella, S. Lessons from emotion psychology for the design of lifelike characters. *Appl. Artif. Intell.* **2005**, *19*, 215–233. [\[CrossRef\]](#)
50. Mulligan, K.; Scherer, K.R. Toward a working definition of emotion. *Emot. Rev.* **2012**, *4*, 345–357. [\[CrossRef\]](#)
51. Scherer, K.R. What are emotions? And how can they be measured? *Soc. Sci. Inf.* **2005**, *44*, 695–729. [\[CrossRef\]](#)
52. Ekman, P.; Friesen, W.V.; Ellsworth, P. *Emotion in the Human Face: Guidelines for Research and an Integration of Findings*; Elsevier: Amsterdam, The Netherlands, 2013.
53. LeDoux, J. The emotional brain, fear, and the amygdala. *Cell. Mol. Neurobiol.* **2003**, *23*, 727–738. [\[CrossRef\]](#) [\[PubMed\]](#)
54. Izard, C.E. The many meanings/aspects of emotion: Definitions, functions, activation, and regulation. *Emot. Rev.* **2010**, *2*, 363–370. [\[CrossRef\]](#)
55. Picard, R.W. Affective Computing for HCI. In *HCI (1)*; Citeseer: Princeton, NJ, USA, 22 August 1999; pp. 829–833.
56. Sundström, P.; Ståhl, A.; Höök, K. In situ informants exploring an emotional mobile messaging system in their everyday practice. *Int. J. Hum. Comput. Stud.* **2007**, *65*, 388–403. [\[CrossRef\]](#)
57. Calvo, R.A.; D'Mello, S.; Gratch, J.M.; Kappas, A. Emotion modeling for social robots. In *The Oxford Handbook of Affective Computing*; Oxford University Press: Oxford, UK, 2014; pp. 296–308.
58. Bartneck, C.; Reichenbach, J.; Breemen, A. In your face, robot! The influence of a character's embodiment on how users perceive its emotional expressions. In *Proceedings of the Design and Emotion 2004*, Ankara, Turkey, 12–14 July 2004.
59. Bates, J. The role of emotion in believable agents. *Commun. ACM* **1994**, *37*, 122–125. [\[CrossRef\]](#)
60. Johnston, O.; Thomas, F. *The Illusion of Life: Disney Animation*; Disney Editions: New York, NY, USA, 1981.
61. Ranjbar, H.; Richards, D.; Bilgin, A.; Kutay, C. First impressions count! The role of the human's emotional state on rapport established with an empathic versus neutral virtual therapist. *IEEE Trans. Affect. Comput.* **2019**. [\[CrossRef\]](#)
62. Feldman, J.M. Four questions about human social behavior. In *Social Psychology and Cultural Context*; SAGE: New York, NY, USA, 1999; pp. 107–113.
63. Moon, Y. *Intimate Self-Disclosure Exchanges: Using Computers to Build Reciprocal Relationships with Consumers*; Division of Research, Harvard Business School: Boston, MA, USA, 1998.
64. Moon, Y.; Nass, C. How “real” are computer personalities? Psychological responses to personality types in human-computer interaction. *Communic. Res.* **1996**, *23*, 651–674. [\[CrossRef\]](#)
65. Moon, Y.; Nass, C. Are computers scapegoats? Attributions of responsibility in human-computer interaction. *Int. J. Hum. Comput. Stud.* **1998**, *49*, 79–94. [\[CrossRef\]](#)
66. Nass, C.; Moon, Y.; Fogg, B.J.; Reeves, B.; Dryer, D.C. Can computer personalities be human personalities. *Int. J. Hum. Comput. Stud.* **1995**, *43*, 223–239. [\[CrossRef\]](#)
67. Isbister, K.; Nass, C. Consistency of personality in interactive characters: Verbal cues, non-verbal cues, and user characteristics. *Int. J. Hum. Comput. Stud.* **2000**, *53*, 251–267. [\[CrossRef\]](#)
68. Ball, G.; Breese, J. *Emotion and Personality in a Conversational Agent*; Embodied Conversational Agents: 2000; MIT Press: Cambridge, MA, USA, 2000.
69. Hanna, N.; Richards, D. The influence of users' personality on the perception of intelligent virtual agents' personality and the trust within a collaborative context. In *International Workshop on Multiagent Foundations of Social Computing*; Springer: Cham, Switzerland, 2015; pp. 31–47.

70. King, W.J.; Ohya, J. The representation of agents: Anthropomorphism, agency, and intelligence. In Proceedings of the Conference Companion on Human Factors in Computing Systems, Vancouver, BC, Canada, 13–18 April 1996; pp. 289–290.
71. Paiva, A.; Dias, J.; Sobral, D.; Aylett, R.; Sobrepepe, P.; Woods, S.; Zoll, C.; Hall, L. Caring for agents and agents that care: Building empathic relations with synthetic agents. In Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems, New York, NY, USA, 19–23 July 2004; Volume 1, pp. 194–201.
72. Berretty, P.M.; Todd, P.M.; Blythe, P.W. Categorization by elimination: A fast and frugal approach to categorization. In Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society, Stanford, CA, USA, 7–10 August 1997; pp. 43–48.
73. Bailenson, J.N.; Yee, N.; Merget, D.; Schroeder, R. The effect of behavioral realism and form realism of real-time avatar faces on verbal disclosure, nonverbal disclosure, emotion recognition, and copresence in dyadic interaction. *Presence Teleoper. Virtual Environ.* **2006**, *15*, 359–372. [[CrossRef](#)]
74. Nowak, K.L.; Biocca, F. The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence Teleoper. Virtual Environ.* **2003**, *12*, 481–494. [[CrossRef](#)]
75. Zajonc, R.B.; Adelman, P.K.; Murphy, S.T.; Niedenthal, P.M. Convergence in the physical appearance of spouses. *Motiv. Emot.* **1987**, *11*, 335–346. [[CrossRef](#)]
76. Aiello, J.R.; Douthitt, E.A. Social facilitation from Triplett to electronic performance monitoring. *Group. Dyn. Theory Res. Pract.* **2001**, *5*, 163. [[CrossRef](#)]
77. Hoyt, C.L.; Blascovich, J.; Swin, K.R. Social inhibition in immersive virtual environments. *Presence* **2003**, *12*, 183–195. [[CrossRef](#)]
78. Catrambone, R.; Stasko, J.; Xiao, J. ECA as user interface paradigm. In *From Brows to Trust*; Springer: Amsterdam, The Netherlands, 2004; pp. 239–267.
79. Davidson, R.; Henderson, R. Electronic performance monitoring: A laboratory investigation of the influence of monitoring and difficulty on task performance, mood state, and self-reported stress levels. *J. Appl. Soc. Psychol.* **2000**, *30*, 906–920. [[CrossRef](#)]
80. Tresselt, M.E.; Mayzner, M.S. Normative solution times for a sample of 134 solution words and 378 associated anagrams. *Psychon. Monogr. Suppl.* **1966**, *1*, 293–298.
81. Rajecki, D.W.; Ickes, W.; Corcoran, C.; Lerner, K. Social facilitation of human performance: Mere presence effects. *J. Soc. Psychol.* **1977**, *102*, 297–310. [[CrossRef](#)]
82. Jackson, J.M.; Williams, K.D. Social loafing on difficult tasks: Working collectively can improve performance. *J. Personal. Soc. Psychol.* **1985**, *49*, 937. [[CrossRef](#)]
83. Beilock, S.L.; Kulp, C.A.; Holt, L.E.; Carr, T.H. More on the fragility of performance: Choking under pressure in mathematical problem solving. *J. Exp. Psychol. Gen.* **2004**, *133*, 584. [[CrossRef](#)]
84. Cahn, J.E. The generation of affect in synthesized speech. *J. Am. Voice I/O Soc.* **1990**, *8*, 1.
85. Bond, C.F.; Titus, L.J. Social facilitation: A meta-analysis of 241 studies. *Psychol. Bull.* **1983**, *94*, 265. [[CrossRef](#)]
86. Erdfelder, E.; Faul, F.; Buchner, A. GPOWER: A general power analysis program. *Behav. Res. Methods Instrum. Comput.* **1996**, *28*, 1–11. [[CrossRef](#)]
87. Jones, E.E. *Ingratiation: A Social Psychological Analysis*; Appleton-Century-Crofts: New York, NY, USA, 1964.
88. Singer, E.; Frankel, M.R.; Glassman, M.B. The effect of interviewer characteristics and expectations on response. *Public Opin. Q.* **1983**, *47*, 68–83. [[CrossRef](#)]
89. Sudman, S.; Bradburn, N.M. *Response Effects in Surveys: A Review and Synthesis*; Aldine Publishing: Chicago, IL, USA, 1974.
90. Nass, C.; Steuer, J. Voices, boxes, and sources of messages: Computers and social actors. *Hum. Commun. Res.* **1993**, *19*, 504–527. [[CrossRef](#)]
91. Oh, C.S.; Bailenson, J.N.; Welch, G.F. A systematic review of social presence: Definition, antecedents, and implications. *Front. Robot. Artif. Intell.* **2018**, *5*, 114. [[CrossRef](#)] [[PubMed](#)]
92. Bente, G.; Rüggenberg, S.; Krämer, N.C.; Eschenburg, F. Avatar-mediated networking: Increasing social presence and interpersonal trust in net-based collaborations. *Hum. Commun. Res.* **2008**, *34*, 287–318. [[CrossRef](#)]
93. Appel, J.; von der Pütten, A.; Krämer, N.C.; Gratch, J. Does humanity matter? Analyzing the importance of social cues and perceived agency of a computer system for the emergence of social reactions during human-computer interaction. *Adv. Hum. Comput. Interact.* **2012**, *2012*, 13. [[CrossRef](#)]
94. Kim, H.; Suh, K.S.; Lee, U.K. Effects of collaborative online shopping on shopping experience through social and relational perspectives. *Inf. Manag.* **2013**, *50*, 169–180. [[CrossRef](#)]
95. Beale, R.; Creed, C. Affective interaction: How emotional agents affect users. *Int. J. Hum. Comput. Stud.* **2009**, *67*, 755–776. [[CrossRef](#)]
96. Brave, S.; Nass, C.; Hutchinson, K. Computers that care: Investigating the effects of orientation of emotion exhibited by an embodied computer agent. *Int. J. Hum. Comput. Stud.* **2005**, *62*, 161–178. [[CrossRef](#)]
97. De Melo, C.M.; Carnevale, P.; Gratch, J. The impact of emotion displays in embodied agents on emergence of cooperation with people. *Presence Teleoper. Virtual Environ.* **2011**, *20*, 449–465. [[CrossRef](#)]
98. Manstead, A.S.; Fischer, A.H. Social appraisal. *Apprais. Process. Emot. Theory Methods Res.* **2001**, 221–232.
99. Van Kleef, G.A. Emotion in conflict and negotiation: Introducing the emotions as social information (easi) model. In *IACM 2007 Meetings Paper 2007*; University of Amsterdam: Amsterdam, The Netherlands, 2007.
100. Adam, H.; Shirako, A.; Maddux, W.W. Cultural variance in the interpersonal effects of anger in negotiations. *Psychol. Sci.* **2010**, *21*, 882–889. [[CrossRef](#)] [[PubMed](#)]

-
101. Adam, H.; Shirako, A. Not all anger is created equal: The impact of the expresser's culture on the social effects of anger in negotiations. *J. Appl. Psychol.* **2013**, *98*, 785. [[CrossRef](#)]
 102. Matsumoto, D.; Yoo, S.H. Toward a new generation of cross-cultural research. *Perspect. Psychol. Sci.* **2006**, *1*, 234–250. [[CrossRef](#)] [[PubMed](#)]
 103. Cohen, D.; Gunz, A. As seen by the other . . . : Perspectives on the self in the memories and emotional perceptions of Easterners and Westerners. *Psychol. Sci.* **2002**, *13*, 55–59. [[CrossRef](#)] [[PubMed](#)]
 104. Wu, S.; Keysar, B. The effect of culture on perspective taking. *Psychol. Sci.* **2007**, *18*, 600–606. [[CrossRef](#)] [[PubMed](#)]
 105. Bickmore, T.W. Relational Agents: Effecting Change Through Human-Computer Relationships. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2003.
 106. Bickmore, T.W.; Puskar, K.; Schlenk, E.A.; Pfeifer, L.M.; Sereika, S.M. Maintaining reality: Relational agents for antipsychotic medication adherence. *Interact. Comput.* **2010**, *22*, 276–288. [[CrossRef](#)]