



Nonverbal communication in virtual reality: Nodding as a social signal in virtual interactions

Nadine Aburumman^{*,a}, Marco Gillies^b, Jamie A. Ward^{b,c}, Antonia F.de C. Hamilton^c

^a Brunel University London, Kingston Ln, London, UK

^b Goldsmiths, University of London, New Cross, London, UK

^c University College London, Bloomsbury, London, UK

ARTICLE INFO

Keywords:

Virtual reality
Social interactions
Nonverbal communication
Head nodding

ABSTRACT

Nonverbal communication is an important part of human communication, including head nodding, eye gaze, proximity and body orientation. Recent research has identified specific patterns of head nodding linked to conversation, namely mimicry of head movements at 600 ms delay and fast nodding when listening. In this paper, we implemented these head nodding behaviour rules in virtual humans, and we tested the impact of these behaviours, and whether they lead to increases in trust and liking towards the virtual humans. We use Virtual Reality technology to simulate a face-to-face conversation, as VR provides a high level of immersiveness and social presence, very similar to face-to-face interaction. We then conducted a study with human-subject participants, where the participants took part in conversations with two virtual humans and then rated the virtual character social characteristics, and completed an evaluation of their implicit trust in the virtual human. Results showed more liking for and more trust in the virtual human whose nodding behaviour was driven by realistic behaviour rules. This supports the psychological models of nodding and advances our ability to build realistic virtual humans.

1. Introduction

Face-to-face interaction is a central part of human life, used to convey ideas, share information, understand intentions and emotions, build trust, and help in learning and decision making. An important goal for researchers in computer science is to engineer virtual systems, including both virtual humans and immersive virtual reality contexts, which can replicate a real face-to-face conversation. It is also an important goal for researchers in psychology to understand how humans behave during interactions, and to test theories about which aspects of these interactions are significant. Whether in a physical or virtual setup, human communication involves both verbal exchanges and nonverbal behaviours. Nonverbal communication is an effective and expressive tool that is employed to send and receive social signals, which humans have been using for thousands of years before they developed the capability to communicate with words. Therefore, both analysis and synthesis of nonverbal communication constitutes an essential part of research in Human-Computer Interaction (HCI) (Garau et al., 2003). Although physical face-to-face communication remains the most

powerful form of interaction, modern communication is frequently mediated by technology and held virtually. Virtual Reality (VR) is a virtual medium of communication that can facilitate the creation of immersive real-time interaction, and enhance social presence in 3D environments (Marini et al., 2012). Moreover, VR allows intrinsic face-to-face communication in a virtual setting and feels verisimilar. In the present study we employ VR in our experiments, as it has an unrivalled potential to impact the future of numerous sectors, including interactive training, teleconferencing, education, consulting, social rehabilitation, tourism, health-care and simulations (Gunkel et al., 2018b; Jack et al., 2001; Lee and Wong, 2008).

Nonverbal communication includes head nodding, eye contact, leaning forward, and body orientation. In particular, head nodding plays an important role in regulating an interaction, signalling who is to take a turn or whether or not someone is interested in and attending to a particular item. This type of signalling is commonly called backchannelling (Cogdill et al., 2001), and is often produced by the listener or follower in a conversation to send subtle messages to the speaker. Hence, including this type of signal in an implementation of a

* Corresponding author.

E-mail address: nadine.aburumman@brunel.ac.uk (N. Aburumman).

URL: <https://nadineab.github.io/webpage/> (N. Aburumman).

<https://doi.org/10.1016/j.ijhcs.2022.102819>

Received 14 March 2021; Received in revised form 19 February 2022; Accepted 21 March 2022

Available online 30 March 2022

1071-5819/© 2022 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

virtual human can be vital in making the human interaction partner feel comfortable and listened to. Previous work in building embodied conversational agents and virtual humans has implemented a variety of nonverbal signals, including nodding (reviewed in Section 2). However, many studies are based on an ad-hoc choice of social signals, while others have not always been replicated. Our starting point is a recent psychological study which carefully characterised the patterns of head movements shown in natural conversation (Hale et al., 2020) and identified two distinct patterns of head nodding.

In this paper, we implemented various experiments involving virtual interaction between a human-controlled avatar (the human participant), and a virtual human whose behaviour is controlled by a computer program. In these experiments, we focused on designing four different nonverbal signalling modes that are highly important in face-to-face human interaction: eye-blinking, nodding, facial expressions, and changing of gaze (see Fig. 1), and we specifically manipulate the nodding behaviour between two different virtual humans. In our experiments we varied the head nodding behaviour of the virtual human in order to measure the effect of nodding on the experience of the human participant. That is, we enabled and disabled the head nodding movements of the virtual human among various runs of the experiments. While the nodding mode is enabled, we modelled two types of head movement based on the findings of Hale et al. (2020): head nod mimicry, in which the virtual human mimics the participant's head nod with a constant 600 ms lag, and fast nods that are executed by the virtual human as a listener in the conversation, by making small rapid head movements. We use VR technology in our virtual social interaction experiments, where we investigate the role of head nodding as visual feedback during conversation. Additionally, we explore how head nodding movements impact the engagement in interaction between the human participant and the virtual human, and we aim to address how it can support virtual communication and promote proximity. The work we present in this paper is thus summarised as follows:

1. We implemented two types of nodding behaviour on a virtual human as social interaction signals in VR, and tested if participants prefer the theory-driven virtual human to a matched virtual human without the nodding behaviour.
2. We implemented a rich version of a novel “virtual maze task” (Hale et al., 2018) in VR, in which participants navigate a maze and have the opportunity to ask for help from different virtual humans. The frequency with which participants approach a virtual human and follow that virtual humans' advice can provide an implicit measure of trust and liking.

2. Related work

Social interaction in Virtual Reality (VR) is a challenging and interdisciplinary area of research that incorporates computer science, human-computer interaction, and psychology. In this section, we review earlier references related to *social signals* and *nonverbal communication*, mainly *head nodding*. We discuss those studies in the literature that are most relevant to our work.

Social signals form the set of nonverbal information that people exchange during social interaction, and they are the basis of highly effective communication. In face-to-face interactions, nonverbal signals make particularly important contributions to the organisation of turn taking and to back-channeling (Gravano and Hirschberg, 2011). Back-channeling is an umbrella term for all the signals which a listener sends to a speaker; these can inform the speaker whether the listener understands and can provide encouragement to the speaker (Lala et al., 2017). Some of the earliest work identifying back-channel signals and exploring their meaning comes from the 1960s (Argyle et al., 1968; Kendon, 1970). More recently, researchers in psychology have focused on mimicry behaviour as an important component of social interaction. Mimicry has been described as a “social glue” (Leander et al., 2012) which causes people to like each other and feel affiliation for each other. In face-to-face interactions, the amount of mimicry (broadly defined) can predict feelings of liking (Salazar Kämpf et al., 2018). However, it is challenging to test the causal role of mimicry during a natural face to face interaction. Building virtual agents who do/do not mimic is an approach by which psychologists can test their theories.

In order to understand back-channeling, mimicry, and other nonverbal types of communication, and in order to explore social signals in conversational dynamics for socially-aware computing, the notions of social signal processing and applied signal processing techniques have been introduced by Pentland (2005, 2007); Pentland et al. (2008). Building on these ideas, there has been a growing body of research aimed at analysing social signals with the goal of building tools, applications, and interfaces for humans, based on models of human behaviour (Andry et al., 2011; Burgoon et al., 2017; Gunes et al., 2008; Vinciarelli et al., 2009). For example, in a study done by Otsuka et al. (2006), a visual head-tracking technique was employed to analyse the interaction between people from audio-visually recorded meetings. Another work, by Byun et al. (2011), analysed nonverbal cues, including head nods and eye gazing of participants in video conferencing, and presented a user study that shows that their predictive framework has the capability to positively influence the conversations in the video conference. An exhaustive survey on past efforts exploring social cues in human behaviour using vision-based, audio-based, and audio-visual



Fig. 1. Our VR experiment, where the participant interacts with a virtual human in a form of structured conversation. Top row: The virtual human is the speaker, sharing and reading information on the tablet. The virtual human performs three nonverbal signals: eye-blinking, facial expressions, and changing of gaze direction. Bottom row: The participant is the speaker and the virtual human executes subtle fast nods as nonverbal feedback.

analysis, can be seen in Vinciarelli et al. (2008) and Zeng et al. (2009).

Over the last couple of decades, *nonverbal communication* of social interaction in *virtual worlds* has been studied extensively, using embodied conversational agents (ECAs) (Cassell et al., 1998; Hömke et al., 2018; Walker et al., 1994). One common objective of these studies is investigating human reactions to different degrees of realism in nonverbal cues given by virtual agents displayed on screen (Kipp, 2005). The paper by Schroder et al. (2012) presented virtual characters with different personalities that induce emotions by producing emotional feedback and nonverbal cues. Furthermore, Lee and Marsella (2006) developed a tool that automates the selection and timing of nonverbal behaviours for virtual agents. In particular, Osugi and Kawahara (2018) studied the effect of *head nodding* and shaking motions on perceptions of likeability, where the authors employed 3D animated characters, and found that the nodding head motion enhanced perceived likeability and approachability, whereas the shaking motion did not influence the ratings. A more recent study on the social impact of people interacting with virtual characters indicates that humans treat virtual characters as if they were real, and exhibit many of the subtle social signals that arise in human-to-human interaction (Miller et al., 2019). These studies suggest that nonverbal behaviours for the virtual agents impact the quality of the social interaction, and this highlights the need for psychological based models to build realistic nonverbal communication for virtual humans.

On the other hand, *Virtual Reality (VR)* is able to provide immersive virtual environments (IVEs) and social presence for users. This technology's availability and affordability has significantly improved over recent years (Mütterlein, 2018; Oh et al., 2019), which makes VR a highly effective medium for communication and social interaction, and has encouraged multiple industrial companies to launch various social VR platforms (e.g., Facebook Spaces and VRChat) (Gunkel et al., 2018a; Tarr et al., 2018). One of the earliest works in social signalling using VR is presented by Bailenson et al. (2003), where the authors ran a set of controlled experiments to determine the impact of gaze behaviour of participants interacting with a virtual agent. The results of their experiments suggest that within VR, proxemic space behaviours and gaze operate almost identical to how they operate in physical face-to-face settings. A number of studies have been carried out that investigate the relationship between mimicry in *head movement* and rapport (Hale and Hamilton, 2016; Hamilton et al., 2016), where a head tracking device was employed to track the participant's head movement, rather than using a head-mounted device (i.e., a VR headset). More recently, (Smith and Neff, 2018) presents a study on verbal and nonverbal communicative behaviour in embodied virtual reality vs. face-to-face interaction, where the results show similar communicative behaviour across these two modes of communication. Furthermore, Maloney et al. (2020) presented observational and interview studies that highlight the nuance of nonverbal communication in social VR compared to offline (face-to-face) and traditional virtual world settings. Sun et al. (2019) studied whether natural occurrence of synchronous nonverbal behaviour in VR correlates with social presence and closeness. Further, Tanenbaum et al. (2020) presented an inventory of nonverbal communication in commercial social VR platforms, where they found that facial expression control, and unconscious body posture remain challenges, and they are the two critical social signals that are currently poorly supported within today's commercial social VR platforms. As noted in many VR experiments and simulations, VR can provide a certain level of control over different sensorimotor, cognitive and social factors, along with the ability to provide co-presence and high interactivity in a virtual environment.

Building on the studies mentioned above, we present a well-controlled study in VR, where we implement realistic nonverbal behaviour, including two types of head movements, i.e., mimicry and fast nodding, on virtual humans. Moreover, we investigate and discuss the impact of head nodding during virtual social interactions, and its influence on the participant's impression of the virtual humans. This

allows us to test psychological theories about the impact of nodding and demonstrate how accurate understanding of natural behaviour can contribute to building virtual humans in immersive environments. Furthermore, based on our previous work and other studies detailed in this section, we can formulate two distinct hypotheses. First, participants will report more liking for and interactivity with the virtual human who shows natural nodding behaviour, as measured on the questionnaires. Second, participants will more often approach the virtual human who showed natural nodding behaviour when they later encounter the virtual human in a maze-solving task.

3. Experiment design

We conducted our experiments in the social interaction lab at University College London, where an ethics approval was secured for the experiment. Participants were recruited to the study and gave written informed consent to take part. We were able to collect data from 21 participants, where there were 15 females and 6 males. The average age of the participants was 24. We used on our VR experiment the HTC VIVE Pro as the head-mounted device, with a display resolution of 1440×1600 pixels per eye. The head and hand movements of the participants are tracked using the HTC VIVE Pro headset and controller. The virtual scenes were created in Unity 18.3, using the SteamVR Application Programming Interface (API).

The appearance of the virtual human can influence the virtual experience and the participant's perception of it. A main point of concern here is the virtual human's visual realism. The study by Mori et al. (2012) presented the *uncanny valley* effect, which describes a negative response from the participants who are interacting with a near photo-realistic virtual human. Also, several studies show that increasing the visual realism of a virtual human's appearance tends to expose the virtual human to a harsher judgment by the participant, which is an important reason for why, generally, realism is considered to be a bad predictor for affinity towards the virtual human (McDonnell et al., 2012; Zell et al., 2015). Therefore, for the appearance and render style of our virtual human, we used a style similar to the CG style described in (McDonnell et al., 2012). Thus, we chose to use a less realistic and cartoon-style, as such stylised virtual humans are often preferred over realistic ones (Zell et al., 2015). Fig. 2 shows the appearance of the virtual humans used in our experiment. We used a within-subjects design for this study, where all participants interact with two virtual humans: One is animated with naturalistic nodding, and the other without. This is because within-subjects designs provide more statistical power and are not impacted by individual differences. All the non-essential features of the stimuli, including the order of interacting with each virtual human, the visual appearance and the voice actors, were counterbalanced across participants so that we can examine just the impact of interactivity in our statistical analysis. In our experiment, participants completed three distinct tasks as follows.

In the first task, the participants were told that they would have a conversation with two different virtual humans in VR, and would be discussing a set of facts about some of the states in the United States of America (for example: the fact that the state sport of Maryland is jousting). In the first stage of the experiment, the participant sits in a



Fig. 2. The virtual humans that we used in our experiments. Anna on the left, and Beth on the right.

virtual office room and has a few minutes to acclimatise to the virtual environment. The participant then watches a short demonstration video of the task as performed by two humans on a virtual TV. This teaches the participants how the trials are structured, with monologue and dialogue segments, and provides a model for how participants should behave, providing some chatter as well as a simple reading of the facts available.

Then, the participant meets the first virtual human (Anna), where Anna introduces herself and prompts the participant to introduce themselves. After that, Anna starts a monologue of 45–55 s, where she reads a set of facts about a US state on a tablet, and shares these facts with the participant. This is followed by a 35–45 s dialogue between Anna and the participant, where the facts are discussed. After that, the roles are swapped: The participant reads facts about a different state for 45–55 s, and the participant and virtual human have a 45–35 s dialogue afterward. See Fig. 3, which illustrates the timeline of for one trial of the task. After a series of four such trials (i.e., consisting of discussing eight different states), the participant meets the second virtual human (Beth). The same procedure is repeated and the participant discusses eight new states with Beth and in each trial two states are discussed. In total, the participant has to complete eight trials, where four are with Anna and four are with Beth. Table 1 contains information about the states that are discussed by the virtual humans and the participants.

The aim of this experiment is to study the effect of head nodding in social interaction. Hence, we designed these two virtual humans (Anna and Beth) to provide identical eye-blinking, facial expressions, and changing of gaze behaviours. The only difference between the behaviour of the two virtual humans is that one virtual human engages in naturalistic nodding behaviour that is contingent on their partner’s actions, while the other shows only pre-recorded head motion. The nodding behaviour of the virtual humans is counterbalanced: For half of the participants, Anna is the one who nods, while Beth does not. For the remaining half of the participants Beth nods, while Anna does not.

The second task used a virtual maze to implicitly measure the proximity, trust, and attraction that the participant has formed towards the virtual humans. In this second task, the participants plays a game where they have to escape a maze. They navigate through the maze by pointing, and teleporting to marked locations in the field-of-view, using the VIVE controllers. The virtual humans Anna and Beth are positioned at eight decision points in the maze where the participant has to choose a direction. The participant can choose to approach either Anna or Beth in order to get advice about how to complete the maze (see Fig. 4). Both virtual humans are programmed to give the same advice (“Go left” / “Go right” / “Go straight ahead”) and the participant would only receive advice from the virtual human that has been approached. Moreover, at each point at which the participant encounters the two virtual humans, the distance between the VR headset (participant) and the two virtual humans’ positions in the virtual space, is the same. We interpret the

Table 1

A series of eight trials, which consists of discussions of 16 different states.

Trial	Virtual human	Who speaks first?	State discussed
1	Anna	Anna (Virtual Human)	Arkansas
2	Anna	Participant	Maryland
3	Anna	Anna (Virtual Human)	Idaho
4	Anna	Participant	Mississippi
5	Anna	Anna (Virtual Human)	Illinois
6	Anna	Participant	Delaware
7	Anna	Anna (Virtual Human)	Oregon
8	Anna	Participant	Colorado
9	Beth	Beth (Virtual Human)	Indiana
10	Beth	Participant	Wyoming
11	Beth	Beth (Virtual Human)	Massachusetts
12	Beth	Participant	Maine
13	Beth	Beth (Virtual Human)	Ohio
14	Beth	Participant	Georgia
15	Beth	Beth (Virtual Human)	Pennsylvania
16	Beth	Participant	Minnesota

number of times the participant approaches each virtual human as a measure of their trust in that virtual human.

Finally, in the third task, the participant removes the VR headset and completes a set of questionnaires, which provides an explicit measure of their attitudes towards each virtual human. The questionnaires all use a seven-point Likert scale, where a score of 1 on the scale represents strong disagreement, and a score of 7 represents strong agreement. The participant rates the experience (i.e., through scoring claims such as “I felt mentally immersed in the experience.”, and “I felt that I experienced a sensation of reality.”). Moreover, the participant rates the two virtual humans, measuring their affinity with them (through, e.g., the claim “I find the virtual human appealing.”) and interactivity (e.g., “I felt that the virtual human held my attention during discussion.” and “I believe the virtual human showed attention for what I was saying.”).

In Virtual Reality, cyber-sickness occurs due to the dissociation between the visual input and the vestibular input. In the design for our experiment, we tried to minimise visual-vestibular conflicts by implementing a teleportation locomotion method in our maze task. This method eliminates the acceleration and/or granular movement allowing the participants to instantly teleport to locations by pointing at them. This style of movement has been shown to reduce cyber-sickness (Coomer et al., 2018). We also adopt dynamic blurring, which reduces cyber-sickness by reducing the amount of visual information and optical flow available to the participant. The participant wears a HTC Vive Pro, where a circular area is in focus which is centered on the participant’s gaze point which is tracked by an eye-tracking device embedded in the headset while the rest of the image is blurred. This strategy was found effective for reducing cyber-sickness (Budhiraja et al., 2017).

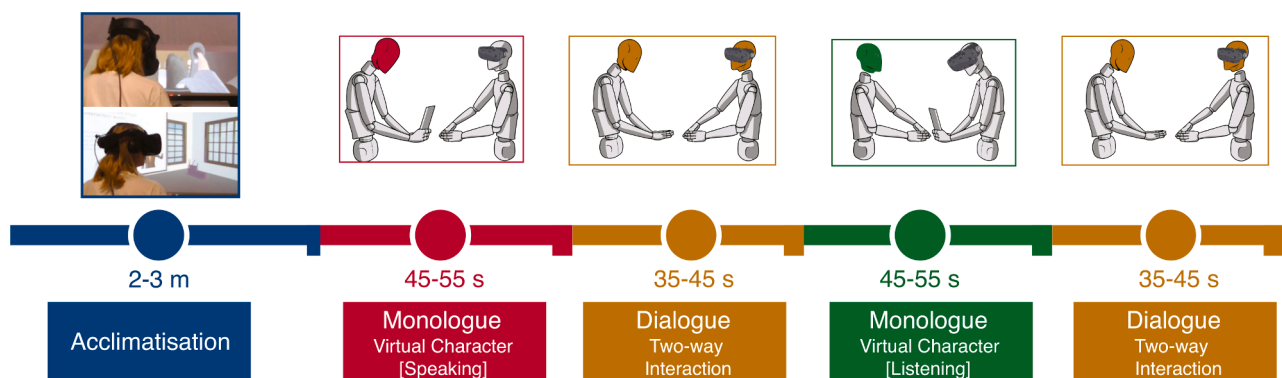


Fig. 3. A timeline for one trial of the first task. The participant acclimatizes to the virtual environment (in blue), then meets the virtual human, where the virtual human starts reading some facts for 45–55 s (*speaking*, in pink). After that, both discuss these facts for 35–45 s (in brown). Subsequently, the participant starts reading for 45–55 s while the virtual human is *listening* (in green), after which, again, they discuss for 35–45 s (in brown). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

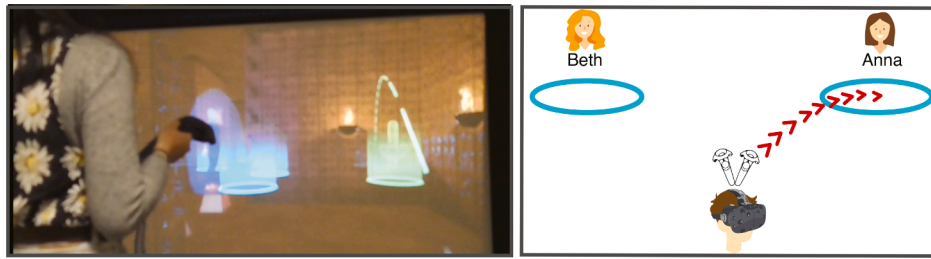


Fig. 4. The second experiment, where the participant has to escape a maze. During this experiment, the participant is able to navigate by pointing to a location in their field-of-view, and teleporting to the pointed location. Anna and Beth are positioned at some of the junctions of the maze, where the participant needs to repeatedly decide from which of the virtual humans to take advice on direction.

4. Experiment implementation

The main purpose of this study is to investigate the role of head nodding as a nonverbal signal in virtual interaction, using VR, in order to facilitate immersive interaction in social communication, remote collaboration, and learning. To this end, we design a VR social experiment where a human participant interacts with two virtual humans (both females).

For the *listening* behaviour, and while the participant is reading and sharing information, we implemented an eye movement module that controls the eye contact and blinking behaviour of the virtual human (see Section 4.1). We also implemented a head movement module, where the virtual human imitates a participant’s head nod if it occurs, and performs fast head nods, i.e., where the head pitch changes at a relatively high frequency. Head nodding movements of the virtual human are programmed to take a fraction of approximately 15–20% of the total speaking time of the participant (see Section 4.3). Furthermore, with regard to the *speaking* behaviour of the virtual human, we make the virtual human perform movements where she changes her gaze direction occasionally: The virtual human interchanges looking at the participant and looking at the tablet from which she is reading. Moreover, we use a lip sync module with a pre-recorded voice of actresses, and we employ a Wizard of Oz module for the choice of facial expressions and for use during the dialogue interaction (see Section 4.2). See Fig. 5 for an overview illustration of the nonverbal modules implemented within our VR experiment.

4.1. Eye movement module

The eye movement module controls the virtual human’s eye-blinking, eye contact, and gaze direction. For the eye-blinking movement, we employ the eye blink model presented by Trutoiu et al. (2011), which is based on physiologically human data. Thus, we use a 9-frame

blink duration model (at 30 frames per second), where the eye blink profile is illustrated in Fig. 6. The spontaneous blink rate (R) of the virtual human is between 6.0 to 30.0 eye blinks per minute: The time between two consecutive blinks is constrained to lie between 2 to 10 s, where after each blink, a new time duration is drawn from the interval of 2 to 10 s, uniformly at random. Moreover, the eye blink rate while the

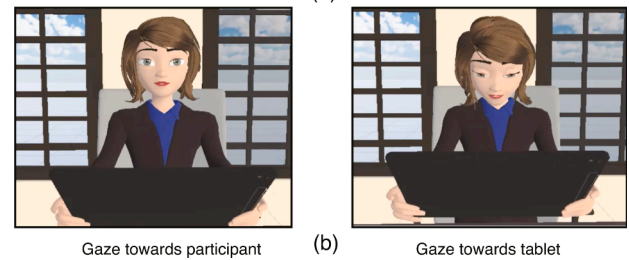
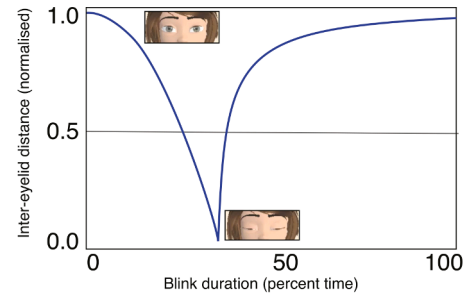


Fig. 6. (a) The eye blink profile used for modelling blinks in our experiment, where a 9-frame blink duration (at 30 fps) model is adopted. (b) The virtual human in our study is sharing information and reading facts, where she interchanges her eye gaze direction between the participant and the tablet.

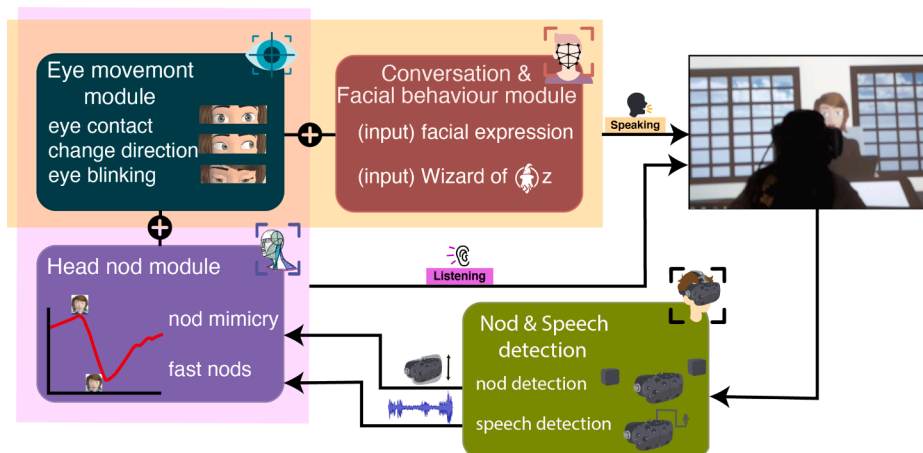


Fig. 5. An overview of the nonverbal modules implemented in our VR experiment. Two modules that control the virtual human’s nonverbal behaviour are active while the virtual human is *listening*: The eye movement module (controlling eye contact and blinking) and the head movement module (controlling mimicry head nodding and fast head nodding). For the virtual human’s *speaking* behaviour, we implemented *change of gaze direction* behaviour within the eye movement module, and we use a *Wizard of Oz* module for the choice of facial expressions and for use during the dialogue interaction.

virtual human is in the reading posture is lower, at 2.4–14.2 eye blinks per minute (Doughty, 2001).

During listening to the participant, the virtual human is always gazing at the participant's headset to indicate eye contact and engagement. On the other hand, while the virtual human is speaking and reading facts from her tablet, and sharing information with the participant, she employs both eye and head movements to replicate human behaviour. The virtual human interchanges her gaze direction between the tablet and the participant, which models a gaze mechanism similar to the one presented by Andrist et al. (2012). The virtual human keeps her head aligned to the participant, when gazing at the participant. When gazing at the tablet, the virtual human aligns her head with the tablet (see Fig. 6). When a change of gaze direction is performed, it is subject to an eye saccade of approximately 15–20°, where the head motion begins approximately 100 ms before the eyes saccade as presented by Lauritis and Robinson (1986).

4.2. Conversation and facial behaviour module

In our experiment, both of the virtual humans' (named Anna and Beth) voices were recorded by native English speaking actresses. We prepared a set of phrases related to each conversation topic (based on a previous study with real participants), including some stock phrases ("yes, tell me more") etc, so that a Wizard of Oz could provide a realistic conversation with the participant. For each prepared speech phrase, we also created a matching facial behaviour in advance. We employ a lip sync algorithm that works off-line (i.e., by preprocessing a recorded voice to compute lip movements), which enables using the voice signal to drive the blendshape deformation of mouth movement and facial expression automatically (Lewis, 1991). In addition, some prepared speech phrases were linked to pre-specified facial expressions, such as a smile with "yes" and a small frown with "really?".

During the dialogue interaction between the virtual human and the participant, we used a human wizard, who interacted using a specific interface within Unity to select an appropriate item from the set of prepared phrases. The human wizard who operates the dialogue was blind to conditions. Because participants will not behave in the same way when in dialogue with the virtual human, we designed the Wizard of Oz module for multimodal utterance to anticipate various dialogue scenarios and facial expressions. This enabled us to create a realistic dialogue for each participant.

4.3. Head nod module

Head nodding plays an important role for information sharing during human interaction. Therefore, in our experiment, we model two types of head movements based on the findings of Hale et al. (2020): mimicry head nodding and fast nodding. The nodding behaviour of the virtual human gives feedback to the participant, and is part of the virtual human's listening behaviour.

The study presented by (Hale et al., 2020) employed motion capture technology to investigate head nodding in the human-human dyadic social interaction (31 dyads), where cross-wavelet analysis was used to examine the data. They found that mimicry occurs between people with a constant lag of around 600 ms. This study suggests that in a conversation, the conversation follower tends to mimic the conversation leader's head nod with roughly a 600 ms delay.

To model mimicry head nodding of the virtual human, we first detect a participant's potential head nod based on movement data of the VR headset. The VR headset utilizes a sensor mounted on the participant's head, which allows for head gesture detection efficiently and accurately. We used the method presented by Katsuomi (2018) to detect head nods based on VR headset movements in real-time. After the participant performs a head nod, with a lag of 600 ms after the detected end of the head nod movement, the virtual human performs a head nod with more or less the same value of the pitch rotation angle (See Fig. 7). The

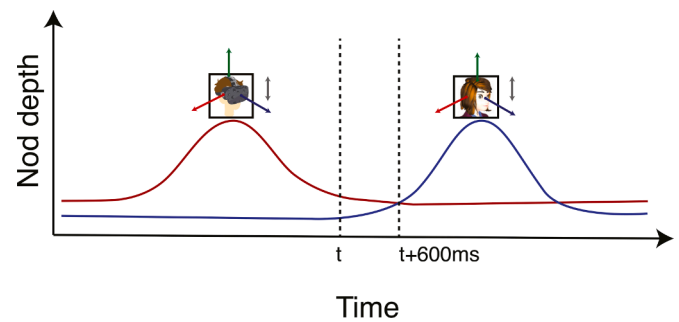


Fig. 7. The virtual human (the follower) mimics the participant's head nod (the leader) with a constant lag of 600 ms and with approximately the same pitch rotation angle (nod depth).

velocity of the head motion lies in the interval $[r \cdot 1.25, r \cdot 6.91]$ and is chosen uniformly at random from that interval, where r is a radius, given by the distance from the center of the virtual human's neck joint to the chin. This nodding motion includes slow-in motion (i.e., motion starts slowly and accelerates), and slow-out motion (i.e., the motion decelerates and then stops).

During the virtual human's listening behaviour and while the participant is speaking, we implemented fast nodding behaviour as visual feedback of agreement towards the participant. Such a fast nod movement of the human consists in 2–3 rapid nods, with faster velocity of the head motion than in the mimicry head nodding behaviour. In this case, the velocity lies in $[r \cdot 16.33, r \cdot 40.84]$. We detect the participant's speech based on the audio coming from the microphone input of the headset, where we filter out noise by using a low-pass filter at a mean cutoff of 10.9 kHz. The virtual human performs these fast head nods in a fraction of approximately 15–20% of the total speaking time of the participant.

For the second virtual human who did not have this nodding module active, the head movement behaviour was done through a "still"-animation captured from very subtle and slow pre-recorded head movements of one pilot participant. This means that the virtual human does not show slow mimicry nods or fast nods, and does not move in a fashion that is contingent on the participant's own movements and speech.

5. Results and discussion

Post-experiment questionnaire To investigate the effect of nodding as a social signal in virtual interaction using VR, we conducted an analysis of the data obtained from 21 participants, where none of the participants have used VR before, and were not familiar with any virtual human. We used Cronbach's alpha coefficient, denoted by α , as it is the most commonly used in sociological research on a seven-point Likert scale (where a score of 1 stands for strong disagreement, and a score of 7 stands for strong agreement). We found that Cronbach's alpha coefficient for the 46 questions in the questionnaire is 0.9306, which indicates high reliability.

For the *immersiveness and co-presence* section of the questionnaire, participants were asked the following questions: Q1.) *I felt mentally immersed in the VR experience.* Q2.) *I felt that I experienced a sensation of reality.* Q3.) *I felt that the experience was natural.* Q4.) *I felt that I was in the presence of another person in the virtual room.* The median rating of Question 1 in this section was 6 out of 7, where the mean M was 5.48 and the standard deviation SD was 1.03. For Question 2, the median rating was 5, where M was 4.77 and SD was 1.27. The median rating for Question 3 was 5, where M was 4.71 and SD was 1.29. Lastly, the median rating for Question 4 of this section was 4, where M was 4.43 and SD was 1.33. See Fig. 8 for an illustration of the result, for this section of the questionnaire.

To examine the effect of head nodding during the virtual social interaction, we carried out t -tests on the ratings given for each of the

Nodding as a Social Signal in Virtual Reality

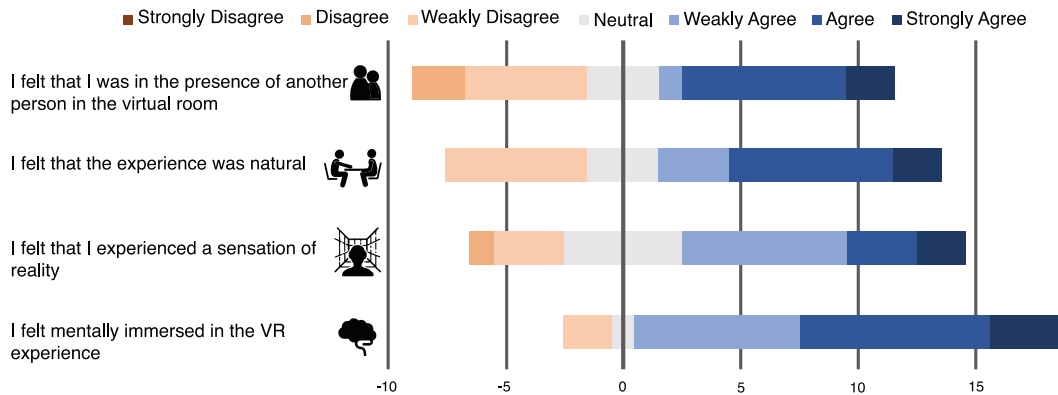


Fig. 8. The results of four questions related to the *immersiveness and co-presence* of the post-experiment questionnaire.

eighteen questions (which include questions on personality, affinity and interactivity) of the post-experiment questionnaire. Those questions are a modified version of the one used in Kidd and Breazeal (2004) and Ho and MacDorman (2010). We found a significant difference in the mean values ($p = 0.036, t = 2.167, df = 40$) for the nodding/social virtual human (mean = 4.92) for the nodding human and the mean was 4.38 for the non-nodding virtual human. The mean rating scores for the virtual humans are represented in Fig. 9.

We also conducted a post-hoc analysis to identify which questions were best at discriminating between the two virtual humans, in order to design future questionnaires. In the interactivity section of the post-experiment questionnaire, the nodding virtual human was rated significantly higher than the non-nodding virtual human, where the mean scores were higher in following questions. Q1.) *I felt that the virtual human was maintaining eye-contact with me*, where the difference in the means (henceforth called the *M-difference*) was 0.905 ($p = 0.0213, t(20) = 2.09$). Q2.) *I felt that the virtual human’s head movement was natural*, where *M-difference* was 1.238 ($p = 0.0003, t(20) = 3.68$). Q3.) *I believe the virtual human showed attention to what I was saying*, where the *M-difference* was 1.24 ($p = 0.0014, t(20) = 3.16$). Therefore, our results indicate that the virtual human’s nodding behaviour while the participant is speaking was perceived as engaging and interactive.

Measuring implicit trust using the maze task The maze task provides an implicit measure of how much participants trust each virtual human. We counted the number of times that participants approach each virtual human (out of 8 decision points). We executed a paired-samples *t-test* to

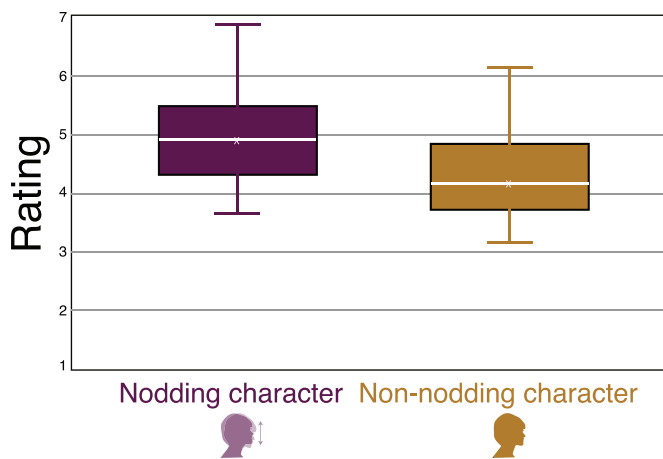


Fig. 9. The difference in the mean rating scores between the nodding virtual human and non-nodding virtual human for eighteen questions on personality, affinity, and interactivity of the post-experiment questionnaire.

determine whether there was a significant difference between the number of approaches towards the nodding virtual human (social virtual human) and non-nodding virtual human (non-social virtual human). Overall, the participants approached the nodding character ($M = 4.62$) more than the non-nodding character ($M = 3.38$), and this result was significant ($p = 0.0453, t(20) = 1.77$). Moreover, the participants always followed the advice from the virtual human that they approached.

As the results in Fig. 10 illustrate, some participants approached the non-nodding virtual human more the nodding virtual human, but overall there were more approaches to the nodding virtual human (purple). This indicates that nodding virtual human is generally been perceived as more approachable and regarded as more trustworthy.

6. Limitations

A limitation of the experiment was that only 21 participants (15 female, 6 male) took part in this experiment. Although the small sample size can decrease the likelihood and inflate the chance of finding statistical significance, that was not the case in this study (Button et al., 2013). However, in our future studies, we will test a higher quantity of

Nodding as a Social Signal in Virtual Reality

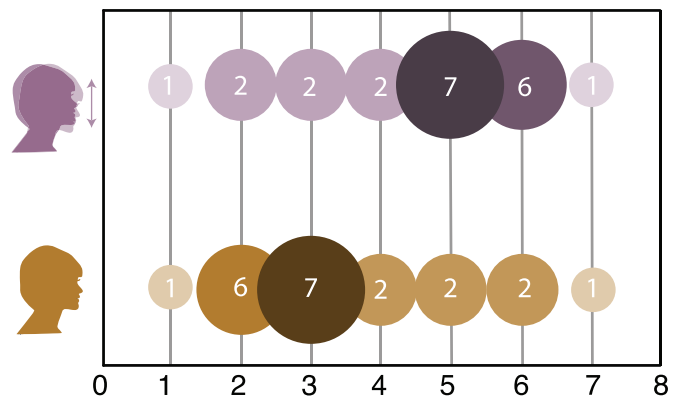


Fig. 10. Summary of the counts with which the participants approached the nodding and non-nodding virtual humans during the maze experiment. The x-axis denotes the count by which a virtual human was approached, the y-axis distinguishes between the nodding (purple) and non-nodding virtual human (brown), and the numbers in the interior show the number of participants who approached the specified virtual human the specified number of times. For example, there were seven participants who, during the maze experiment, approached the nodding virtual human 5 times (and the non-nodding virtual human 3 times). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

participants (over 70 participants) to investigate whether nonverbal and subtle signals contribute to learning, and the quality of peer collaboration within educational settings using VR.

Another point worth mentioning here is that the pre-recorded movements that were used for the “quiet sitting” animation of the virtual character are captured from a single pilot participant. While we paid attention to selecting these movements to be very small, slow, natural movements, that do not appear odd, having a larger sample for these movements available would be useful to ascertain this point somewhat more confidently.

7. Conclusion and future work

In this study, we investigated the role of head nodding behaviour during virtual social interaction using Virtual Reality (VR). We aimed to demonstrate the capacity of virtual humans to embody interactive head nodding behaviour based closely on natural movements, and to test the psychological theory that mimicry of nodding leads to liking and affiliation. Thus, we implemented two sophisticated virtual humans who engage in a simple Wizard of Oz assisted conversation, supplemented by nonverbal signals programmed in VR, including eye-blinking, facial expressions, and natural nodding. We find that participants rate the nodding virtual human more positively in our explicit questionnaire measures, and also approach the nodding virtual human more often than the non-nodding virtual human in our maze task which measures implicit trust. We consider here what these results mean for the improvements of designing engaging virtual human characters in VR, where all the measures used in our experiment is reported in Section 5.

Previous studies of nodding mimicry using virtual humans have found mixed results, with both positive (Bailenson and Yee, 2005) and negative (Hale and Hamilton, 2016) findings. In our experiments, we used more advanced virtual humans presented in VR, whose nodding behaviour was carefully specified based on findings of (Hale et al., 2020) for natural human movements. The two types of head nodding that we have implemented are (a.) head nod mimicry, in which the virtual human mimics the participant’s head nod with a constant 600 ms lag, and (b.) fast nods, which occur when the participant is speaking. We find a positive impact of naturalistic nodding, with both the implicit and explicit measures showing that participants liked and trusted the nodding virtual human: The nodding virtual human was rated significantly higher than the non-nodding virtual human during the virtual social interaction. The participants have rated particularly highly the question *I believe the virtual human showed attention to what I was saying* of the nodding virtual human, which suggests that the nodding behaviour of the virtual human, while the participant is speaking, was perceived as a sign of engagement and interactivity. Moreover, we found in the maze game experiment that the participants have approached the nodding virtual human for advice more often than they approached the non-nodding virtual human, which appears to reflect higher levels of approachability, liking, or trust towards the nodding virtual human.

This supports the claim that mimicry is a social glue (Lakin et al., 2003), and that by copying the actions of another person, it is possible to build trust and liking. Future studies could test how this generalises across different types of conversation and across different social groups, for example, using virtual humans of different races and genders. Our results also demonstrate the potential of implementing simple behaviour rules derived from real behaviour in virtual humans and virtual reality.

One potential application of this kind of research is in generating artificial agents who can teach participants new information and can act as a tutor who builds trust. The ‘talking about US states’ task which we used here is well-suited to learning studies because it includes a set of discrete facts which are not known to our (UK) participants before the study begins, but which can be tested afterwards. In our future experiments, we plan to study the effects of nonverbal signs on participants’ performance within VR educational settings. We aim at exploring whether nonverbal and subtle signals contribute to the learning process,

and the quality of peer collaboration and communication in VR. We will, moreover, design more elaborate sets of experiments to investigate to what extent, and in which settings, nonverbal communication adds value to social interaction, in comparison to purely verbal communication, and how it can be used to convey ideas and meaning better than using words only, in order to use VR in educational settings and in academic video conferencing.

Credit author statement

N.A. designed the study, developed the experiment, led the data collection, data analyses and manuscript preparation; A.H. contributed to the study design and data analyses; J.A.W. contributed to the study design; M.G. assisted with technical issues with the development; A.H. and M.G. supervised all aspects of the experiment and manuscript preparation. A.H. also coordinated the research project.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: Antonia F de C Hamilton acknowledges financial support from the Leverhulme Trust under the grant code RPG-2016-251.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at [10.1016/j.ijhcs.2022.102819](https://doi.org/10.1016/j.ijhcs.2022.102819)

References

- Andrist, S., Pejisa, T., Mutlu, B., Gleicher, M., 2012. Designing effective gaze mechanisms for virtual agents. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, pp. 705–714. <https://doi.org/10.1145/2207676.2207777>.
- Andry, P., Blanchard, A., Gaussier, P., 2011. Using the rhythm of nonverbal human - robot interaction as a signal for learning. *IEEE Trans. Auton. Ment. Dev.* 3 (1), 30–42.
- Argyle, M., Lalljee, M., Cook, M., 1968. The effects of visibility on interaction in a dyad. *Hum. Relat.* 21 (1), 3–17.
- Bailenson, J.N., Blascovich, J., Beall, A.C., Loomis, J.M., 2003. Interpersonal distance in immersive virtual environments. *Personal. Social Psychol. Bull.* 29 (7), 819–833. <https://doi.org/10.1146/167203029007002>. PMID: 15018671
- Bailenson, J.N., Yee, N., 2005. Digital chameleons: automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychol. Sci.* 16 (10), 814–819.
- Budhiraja, P., Miller, M. R., Modi, A. K., Forsyth, D. A., 2017. Rotation blurring: use of artificial blurring to reduce cybersickness in virtual reality first person shooters. *CoRR abs/1710.02599*<http://arxiv.org/abs/1710.02599>.
- Burgoon, J.K., Magnenat-Thalmann, N., Pantic, M., Vinciarelli, A., 2017. *Social Signal Processing*. Cambridge University Press. <https://books.google.co.uk/books?id=Md7FDgAAQBAJ>
- Button, K.S., Ioannidis, J., Mokrysz, C., Nosek, B.A., Flint, J., Robinson, E.S.J., Munafo, M.R., 2013. Power failure: why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci.* 14 (5), 365–376.
- Byun, B., Awasthi, A., Chou, P.A., Kapoor, A., Lee, B., Czerwinski, M., 2011. Honest signals in video conferencing. Proceedings of the 2011 IEEE International Conference on Multimedia and Expo. IEEE Computer Society, USA, pp. 1–6. <https://doi.org/10.1109/ICME.2011.6011855>.
- Cassell, J., Pelachaud, C., Badler, N., Steedman, M., Achorn, B., Becket, T., Douville, B., Prevost, S., Stone, M., 1998. *Animated Conversation: Rule-Based Generation of Facial Expression, Gesture and Spoken Intonation for Multiple Conversational Agents*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, pp. 582–591.
- Cogdill, S., Fanderclai, T.L., Kilborn, J., Williams, M.G., 2001. Backchannel: whispering in digital conversation. Proceedings of the 34th Annual Hawaii International Conference on System Sciences, p. 8.
- Coomer, N., Bullard, S., Clinton, W., Williams-Sanders, B., 2018. Evaluating the effects of four VR locomotion methods: Joystick, arm-cycling, point-tugging, and teleporting. Proceedings of the 15th ACM Symposium on Applied Perception. Association for

- Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/3225153.3225175>.
- Doughty, M.J., 2001. Consideration of three types of spontaneous eyeblink activity in normal humans: during reading and video display terminal use, in primary gaze, and while in conversation. *Optom. Vis. Sci.* 78 (10), 712–725.
- Garau, M., Slater, M., Vinayagamoorthy, V., Brogni, A., Steed, A., Sasse, M.A., 2003. The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, pp. 529–536. <https://doi.org/10.1145/642611.642703>.
- Gravano, A., Hirschberg, J., 2011. Turn-taking cues in task-oriented dialogue. *Comput. Speech Lang.* 25 (3), 601–634. <https://doi.org/10.1016/j.csl.2010.10.003>.
- Gunes, H., Piccardi, M., Pantic, M., 2008. From the lab to the real world: affect recognition using multiple cues and modalities. In: Or, J. (Ed.), *Affective Computing*. IntechOpen, Rijeka. <https://doi.org/10.5772/6180>.
- Gunkel, S., Stokking, H., Prins, M., Niamut, O., Siahaan, E., Cesar, P., 2018. Experiencing virtual reality together: social VR use case study. Proceedings of the 2018 ACM International Conference on Interactive Experiences for TV and Online Video. Association for Computing Machinery, New York, NY, USA, p. 233238. <https://doi.org/10.1145/3210825.3213566>.
- Gunkel, S.N.B., Stokking, H.M., Prins, M.J., van der Stap, N., Haar, F.B.T., Niamut, O.A., 2018. Virtual reality conferencing: multi-user immersive VR experiences on the web. Proceedings of the 9th ACM Multimedia Systems Conference, pp. 498–501.
- Hale, J., Hamilton, A., 2016. Testing the relationship between mimicry, trust and rapport in virtual reality conversations. *Sci. Rep.* 6 (1), 1–11. <https://doi.org/10.1038/srep35295>.
- Hale, J., Payne, M.E.M., Taylor, K.M., Paoletti, D., de C. Hamilton, A.F., 2018. The virtual maze: a behavioural tool for measuring trust. *Q. J. Exp. Psychol.* 71 (4), 989–1008. <https://doi.org/10.1080/17470218.2017.1307865>. PMID: 28303739
- Hale, J., Ward, J.A., Buccheri, F., Oliver, D., de C. Hamilton, A.F., 2020. Are you on my wavelength? Interpersonal coordination in dyadic conversations. *J. Nonverbal Behav.* 44 (1), 63–83.
- Hamilton, A., Pan, X.S., Forbes, P., Hale, J., 2016. Using virtual characters to study human social cognition. In: Traum, D., Swartout, W., Khooshabeh, P., Kopp, S., Scherer, S., Leuski, A. (Eds.), *Intelligent Virtual Agents*. Springer International Publishing, Cham, pp. 494–499.
- Ho, C.-C., MacDorman, K.F., 2010. Revisiting the uncanny valley theory: developing and validating an alternative to the godspeed indices. *Comput. Hum. Behav.* 26 (6), 1508–1518. <https://doi.org/10.1016/j.chb.2010.05.015>. Online Interactivity: Role of Technology in Behavior Change
- Hömke, P., Holler, J., Levinson, S.C., 2018. Eye blinks are perceived as communicative signals in human face-to-face interaction. *PLoS One* 13 (12), e0208030.
- Jack, D., Boian, R., Merians, A.S., Tremaine, M., Burdea, G.C., Adamovich, S.V., Recce, M., Poizner, H., 2001. Virtual reality-enhanced stroke rehabilitation. *IEEE Trans. Neural Syst. Rehabil. Eng.* 9 (3), 308–318.
- Katsuomi, K., 2018. VR Gesture Recognizer [Source code]. <https://github.com/korinV/RVGestureRecognizer>
- Kendon, A., 1970. Movement coordination in social interaction: some examples described. *Acta Psychol.* 32, 101–125.
- Kidd, C.D., Breazeal, C., 2004. Effect of a robot on user perceptions. 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566), vol. 4, pp. 3559–3564. <https://doi.org/10.1109/IROS.2004.1389967>.
- Kipp, M., 2005. *Gesture Generation by Imitation: From Human Behavior to Computer Character Animation*. Universal-Publishers.
- Lakin, J.L., Jefferis, V.E., Cheng, C.M., Chartrand, T.L., 2003. The chameleon effect as social glue: evidence for the evolutionary significance of nonconscious mimicry. *J. Nonverbal Behav.* 27 (3), 145–162.
- Lala, D., Milhorat, P., Inoue, K., Ishida, M., Takanashi, K., Kawahara, T., 2017. Attentive listening system with backchanneling, response generation and flexible turn-taking. Proceedings of the 18th Annual SIGDial Meeting on Discourse and Dialogue. Association for Computational Linguistics, Saarbrücken, Germany, pp. 127–136. <https://doi.org/10.18653/v1/W17-5516>.
- Lauritis, V.P., Robinson, D.A., 1986. The vestibulo-ocular reflex during human saccadic eye movements. *J. Physiol.* 373 (1), 209–233.
- Leander, N.P., Chartrand, T.L., Bargh, J.A., 2012. You give me the chills: embodied reactions to inappropriate amounts of behavioral mimicry. *Psychol. Sci.* 23 (7), 772–779.
- Lee, E.A.-L., Wong, K.W., 2008. *A Review of Using Virtual Reality for Learning*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 231–241.
- Lee, J., Marsella, S., 2006. Nonverbal behavior generator for embodied conversational agents. Proceedings of the 6th International Conference on Intelligent Virtual Agents. Springer-Verlag, Berlin, Heidelberg, p. 243255. https://doi.org/10.1007/11821830_20.
- Lewis, J., 1991. Automated lip-sync: background and techniques. *J. Vis. Comput. Animat.* 2 (4), 118–122. <https://doi.org/10.1002/vis.4340020404>.
- Maloney, D., Freeman, G., Wohn, D.Y., 2020. “Talking without a voice”: understanding non-verbal communication in social virtual reality. *Proc. ACM Hum.-Comput. Interact.* 4 (CSCW2) <https://doi.org/10.1145/3415246>.
- Marini, D., Folgieri, R., Gadia, D., Rizzi, A., 2012. Virtual reality as a communication process. *Virtual Real.* 16 (3), 233–241.
- McDonnell, R., Bredt, M., Bühlhoff, H.H., 2012. Render me real? Investigating the effect of render style on the perception of animated virtual humans. *ACM Trans. Graph.* 31 (4) <https://doi.org/10.1145/2185520.2185587>.
- Miller, M.R., Jun, H., Herrera, F., Yu Villa, J., Welch, G., Bailenson, J.N., 2019. Social interaction in augmented reality. *PLoS One* 14 (5), e0216290.
- Mori, M., MacDorman, K.F., Kageki, N., 2012. The uncanny valley [from the field]. *IEEE Robot. Autom. Mag.* 19 (2), 98–100.
- Mütterlein, J., 2018. The three pillars of virtual reality? Investigating the roles of immersion, presence, and interactivity. Proceedings of the 51st Hawaii International Conference on System Sciences.
- Oh, C., Herrera, F., Bailenson, J., 2019. The effects of immersion and real-world distractions on virtual social interactions. *Cyberpsychol., Behav., Social Netw.* 22 (6), 365–372. <https://doi.org/10.1089/cyber.2018.0404>. PMID: 31188686
- Osugi, T., Kawahara, J.I., 2018. Effects of head nodding and shaking motions on perceptions of likeability and approachability. *Perception* 47 (1), 16–29. <https://doi.org/10.1177/0301006617733209>.
- Otsuka, K., Yamato, J., Takemae, Y., Murase, H., 2006. Conversation scene analysis with dynamic Bayesian network Based on visual head tracking. 2006 IEEE International Conference on Multimedia and Expo, pp. 949–952.
- Pentland, A., 2005. Socially aware, computation and communication. *Computer* 38 (3), 33–40.
- Pentland, A., 2007. Social signal processing. *IEEE Signal Process. Mag.* 24 (4), 108–111. <https://doi.org/10.1109/MSP.2007.4286569>.
- Pentland, A., Pentland, S., Heibeck, T., 2008. *Honest Signals: How They Shape Our World*. Bradford Books, Milrazones. <https://books.google.co.uk/books?id=Zt88mwEACAAJ>
- Salazar Kämpf, M., Liebermann, H., Kerschreiter, R., Krause, S., Nestler, S., Schmulke, S. C., 2018. Disentangling the sources of mimicry: social relations analyses of the link between mimicry and liking. *Psychol. Sci.* 29 (1), 131–138.
- Schroder, M., Bevacqua, E., Cowie, R., Eyben, F., Gunes, H., Heylen, D., ter Maat, M., McKeown, G., Pammi, S., Pantic, M., Pelachaud, C., Schuller, B., de Sevin, E., Valstar, M., Wollmer, M., 2012. Building autonomous sensitive artificial listeners. *IEEE Trans. Affect. Comput.* 3 (2), 165–183.
- Smith, H.J., Neff, M., 2018. Communication behavior in embodied virtual reality. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, pp. 1–12. <https://doi.org/10.1145/3173574.3173863>.
- Sun, Y., Shaikh, O., Won, A.S., 2019. Nonverbal synchrony in virtual reality. *PLoS One* 14 (9), e0221803.
- Tanenbaum, T.J., Hartoonian, N., Bryan, J., 2020. “How do i make this thing smile?”: an inventory of expressive nonverbal communication in commercial social virtual reality platforms. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, pp. 1–13. <https://doi.org/10.1145/3313831.3376606>.
- Tarr, B., Slater, M., Cohen, E., 2018. Synchrony and social connection in immersive virtual reality. *Sci. Rep.* 8 (1), 1–8.
- Trutou, L.C., Carter, E.J., Matthews, I., Hodgins, J.K., 2011. Modeling and animating eye blinks. *ACM Trans. Appl. Percept.* 8 (3) <https://doi.org/10.1145/2010325.2010327>.
- Vinciarelli, A., Pantic, M., Bourlard, H., 2009. Social signal processing: survey of an emerging domain. *Image Vis. Comput.* 27 (12), 1743–1759. <https://doi.org/10.1016/j.imavis.2008.11.007>. Visual and multimodal analysis of human spontaneous behaviour
- Vinciarelli, A., Pantic, M., Bourlard, H., Pentland, A., 2008. Social Signal Processing: State-of-the-Art and Future Perspectives of an Emerging Domain. Association for Computing Machinery, New York, NY, USA, pp. 1061–1070. <https://doi.org/10.1145/1459359.1459573>.
- Walker, J.H., Sproull, L., Subramani, R., 1994. Using a human face in an interface. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, pp. 85–91. <https://doi.org/10.1145/191666.191708>.
- Zell, E., Aliaga, C., Jarabo, A., Zibrek, K., Gutierrez, D., McDonnell, R., Botsch, M., 2015. To stylize or not to stylize? The effect of shape and material stylization on the perception of computer-generated faces. *ACM Trans. Graph.* 34 (6) <https://doi.org/10.1145/2816795.2818126>.
- Zeng, Z., Pantic, M., Roisman, G.I., Huang, T.S., 2009. A survey of affect recognition methods: audio, visual, and spontaneous expressions. *IEEE Trans. Pattern Anal. Mach. Intell.* 31 (1), 39–58.