



Review When Agents Become Partners: A Review of the Role the Implicit Plays in the Interaction with Artificial Social Agents

Sanobar Dar * Dand Ulysses Bernardet

College of Engineering and Physical Sciences, Aston University, Birmingham B4 7ET, UK; u.bernardet@aston.ac.uk

* Correspondence: darsf@aston.ac.uk

Received: 30 October 2020; Accepted: 19 November 2020; Published: 22 November 2020



Abstract: The way we interact with computers has significantly changed over recent decades. However, interaction with computers still falls behind human to human interaction in terms of seamlessness, effortlessness, and satisfaction. We argue that simultaneously using verbal, nonverbal, explicit, implicit, intentional, and unintentional communication channels addresses these three aspects of the interaction process. To better understand what has been done in the field of Human Computer Interaction (HCI) in terms of incorporating the type channels mentioned above, we reviewed the literature on implicit nonverbal interaction with a specific emphasis on the interaction between humans on the one side, and robot and virtual humans on the other side. These Artificial Social Agents (ASA) are increasingly used as advanced tools for solving not only physical but also social tasks. In the literature review, we identify domains of interaction between humans and artificial social agents that have shown exponential growth over the years. The review highlights the value of incorporating implicit interaction capabilities in Human Agent Interaction (HAI) which we believe will lead to satisfying human and artificial social agent team performance. We conclude the article by presenting a case study of a system that harnesses subtle nonverbal, implicit interaction to increase the state of relaxation in users. This "Virtual Human Breathing Relaxation System" works on the principle of physiological synchronisation between a human and a virtual, computer-generated human. The active entrainment concept behind the relaxation system is generic and can be applied to other human agent interaction domains of implicit physiology-based interaction.

Keywords: verbal and nonverbal communication; interaction; explicit and implicit interaction; human to human interaction; human to human communication; human-robot interaction; human-agent interaction; human virtual-human interaction; gaze; proxemics; physiological synchrony; entrainment

1. Introduction

In this narrative review, we survey research on implicit interaction between humans and computers—more specifically, the interaction between humans and Artificial Social Agents. Our goal is to better understand how implicit cues can be used to build natural and satisfactory human computer interaction systems and to identify gaps in the research. One of the areas which we identify as under-researched is implicit physiology-driven interaction. To further this area, we present a framework that uses implicit physiological interaction between a human and an Artificial Social Agent. The framework harnesses the phenomenon of physiological entrainment found in human to human interaction. The results from our pilot study support the notion that subtle physiological cues can be used as implicit signals in a HCI system.

Since the inception of computers, the way we interact with them has come a long way, both in terms of hardware and user interface; computers have evolved from bulky static desktops to lightweight tablets, and from text-based system to gesture and voice-based assistants. Griffin reports that out of the global population, 27% use voice searches on mobile and that the number of voice searches doubled from 2015 to 2016 [1]. This shows that voice-based human computer interactions are increasingly becoming a part of daily lives. We use them, e.g., for ordering food and having them read the news to us. This high adoption might be because speech is an efficient form of data input—humans can speak 150 words per minute compared to 40 words when typing [2]—and because speaking does not require learning new skills. Voice-based assistive systems are a good example of the human-computer interaction paradigm that works well for query response interactions such as web searches, but that fail and lead to frustrated users as soon as the interaction becomes more complex [3]. In a study by TheManifest, 95% of users of voice-based assistive systems experienced frustration, where the system either misunderstood the query or was triggered accidentally [4]. It is fair to say that, despite recent progress, interaction with computers does not feel completely natural and there is consensus that HCI is still not on par with human to human interaction [5,6].

After reviewing the literature on mediated and direct human to human interaction, we identified three aspects that facilitate natural, i.e., human to human-like interaction between humans and computers: seamlessness, effort, and satisfaction.

We will conclude this article with a case study where we present a novel system based on multichannel interaction between a human and an Artificial Social Agent. It explores the use of nonverbal, implicit, unintentional interaction with an agent. Concretely, we are describing a system that harnesses the established mechanisms of physiological synchrony, found in human to human interaction, between human and an artificial human, generated in immersive virtual reality. Our system sets the stage for future research in nonverbal and implicit interaction capabilities in HAI.

2. The Process of Interaction

Human to human interaction is a seamless, effortless, and satisfactory process. Seamlessness characterizes a bi-directional interaction that happens continuously without interruptions or delays in communication between the interacting parties. For example, human to human interaction in the real physical world is more seamless than an exchange using a video-calling medium; video calling, in turn, is more seamless than talking over a telephone, and speaking on a telephone is more seamless than using a text-based interface. The importance of seamlessness is further supported by research suggesting that interaction delays on conference systems can have a negative impact on users [7]. Besides, delays can lead to people perceiving the responder as less friendly or focused [8]. Doing any kind of task involves the use of mental resources and requires a certain "effort". Interacting with others, for instance, requires an effort in order to send and receive information. Low effort tasks typically require less attention, are easier to switch between and can be done for a longer period. Studies have shown that face-to-face interaction takes less effort and is less tiring than a video chat [9].

Ultimately, most interactions either with humans or with computers have a purpose. Hence, for an interaction to be deemed "satisfactory", the purpose of the interaction should be achieved with minimal effort and in a minimum amount of time. Voice-base assistants are an example where users spend more time communicating the same information to a computer as compared to another human as they tend to talk slowly and precisely to avoid the need for repetition. Though the purpose of the interaction might eventually be fulfilled, the time taken is longer, which in turn leads to an inefficient interaction and in many cases, the user perceiving the interaction as less satisfactory.

2.1. Dimensions of Interaction

"Communication" refers to a process where information is sent, received, and interpreted. If this information exchange results in a reciprocal influence, i.e., a back and forth between sender and receiver, we refer to it as an "interaction". Such a process of interaction between the two agents is illustrated in

Figure 1. The arrows indicate the transfer of signals between interaction partners. For example, in a negotiation, people interact to discuss and agree on joint action. One partner discloses their viewpoint, and the other partner responds with their viewpoint. This process of interchangeable responses and actions eventually leads to the goal of the process which is to reach a compromise that is agreeable to all. This example shows that interaction is a socio-cognitive process in which we not only consider the other person but also the surroundings and social situations.

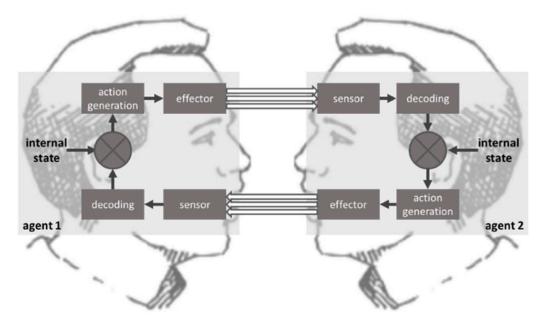


Figure 1. The process of interaction between the two agents. Information is sensed and decoded. In conjunction with the internal state, received messages generate actions through the effector block.

An important aspect of interaction is the means of information transfer from sender to receiver. Rather than over a single, sequential channel, information is often transmitted using multiple channels at once (as illustrated by the multiple arrows between the two agents in Figure 1). In multi-channel communication, stimuli are simultaneously transmitted through different sensory modalities such as sight, sound, and touch. Characteristically, the individual modalities are used for transmitting more than a single message. In the auditory domain, the sender can simultaneously transmit language-based verbal information with prosodic information in the form of the tone of voice. Similarly, the visual domain can be used to transmit language-based messages in the form of sign language and lip-reading. The visual modality affords a broad range of nonverbal channels in the form of eye gaze and facial expressions; people can make thousands of different facial expressions along with varied eye gaze, and each can communicate a different cue. Smiling, frowning, blinking, squinting, eyebrow twitch or nostril flare are some of the most salient facial expressions. Beyond the face, other sources of kinesics information include gestures, head nods, posture, and physical distance. Often people "talk using their hands" emphasizing the use of hand gestures while communicating. How people sit, stand, walk, etc., gives vital cues about how they are perceived by their interaction partners.

While the role of olfactory channels in humans is subtle and still debated [10], haptic channels play an important role in interpersonal communication where a simple handshake is used to communicate agreement. In the above example of negotiation, the simultaneous transfer of both, verbal and nonverbal cues plays an important role; feelings of nervousness or hesitancy expressed through body language or lack of eye contact are perceived as insincerity and disagreement [11].

Broadly speaking, interactions between humans can be classified along the three dimensions of verbal–nonverbal, intentional–unintentional and implicit–explicit (Figure 2). We have already discussed the dichotomy of verbal vs. nonverbal communication. While in an intentional communicative act the sender is aware of what they want to transmit, unintentional signals are sent without a conscious

decision by the sender. Deitic movements to indicate direction are a typical example for the former, while blushing is an example of the latter.

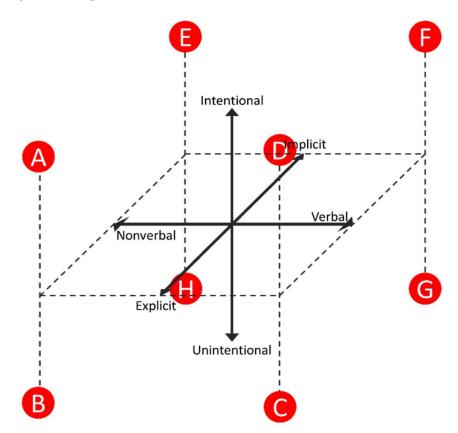


Figure 2. Communication can be classified along the three dimensions of verbal–nonverbal, implicit–explicit, and intentional–unintentional. Head nods during an interaction are an example of nonverbal, explicit, and unintentional communication (**A**). "Freudian slips" fall in the unintentional, verbal and explicit quadrant (**C**). An instruction such as "bring me a glass of water" are verbal-explicit-intentional (**D**). Irony is an example of communication that combines verbal-implicit-intentional (**F**) with simultaneous nonverbal–implicit–intentional communication (**E**). Nonverbal–implicit–unintentional cues (**H**) are closely related to emotion and, therefore, associated with physiological reactions such as sweating and blushing. While some of the examples are straightforward, others seem to be more complicated, e.g., expression of anger or sadness are often explicit, verbal/nonverbal, and unintentional (**G**,**B**).

The third dimension of implicit vs. explicit is defined by the degree to which symbols and signs are used. Explicit interaction is a process that uses an agreed system of symbols and signs, the prime example being the use of spoken language and emblematic gestures. Conversely, implicit interaction is where the content of the information is suggested independent of the agreed system of symbols and signs [12,13]. In an implicit interaction, the sender and the receiver infer each other's communicative intent from the behaviour.

2.2. The Role of Implicit Interaction

During interaction, the reciprocal influence of interaction partners can result in tight coupling of behaviours. An example is rhythmic clapping, where individuals readily synchronize their behaviour to match stimuli, i.e., to match the neighbour's claps that precede their clap [14]. A similar, implicit and unintentional form of reciprocal interaction is the synchronizing of physiological states, where both interaction partners fall into a rhythm. This phenomenon of the interdependence of physical activity in interaction partners is known as physiological synchrony [15]. Studies have shown that romantic partners exhibit physiological synchrony in the form of heart rate adaptation [16,17]. Mothers and their

infants coordinate their heart rhythms during social interactions. This visuo-affective social synchrony is believed to have a direct effect on the attachment process [18]. In psychology, the phenomenon where interaction partners unintentionally and unconsciously mimic each other's speech patterns, facial expressions, emotions, moods, postures, gestures, mannerisms, and idiosyncratic movements is known as the "Chameleon Effect" [19]. Interestingly, this mimicking has the effect of increasing the sense of rapport and liking between the interaction partners [20].

2.3. Interaction Configurations

Interactions are taking place in what we can refer to as different configurations (Figure 3). The distinguishing feature of the configurations is the number of channels and the use of implicit/explicit, verbal/nonverbal and intentional/unintentional messages.

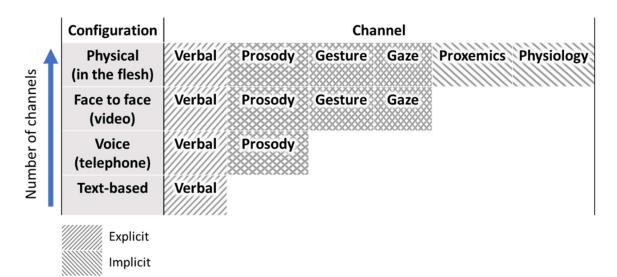


Figure 3. Overview of the relationship between communication configuration, number of channels, and channel properties. The number of channels decreases from physical world interaction to video-calling to voice-calling to texting.

The illustration makes it clear why physical interaction trumps text-based interaction: Multiple parallel channels reduce the risk of interruptions, increasing the seamlessness of the interaction [21]. The redundancy afforded by multichannel communication increases the overall reliability of the communication, and because the decoding is distributed over a larger number of modalities and decoding modules, e.g., for explicit and implicit messages, the overall effort is reduced, even though in physical interaction there is a greater number of channels to decode [22–24]. Ultimately, the efficiency of the communication and the ability to keep track of the goal achievement progress leads to a more satisfactory interaction. This view is supported by research suggesting that incorporating implicit cues in HCI is important for an unambiguous and clear interaction [25]. Furthermore, multiple channels lead to quicker task completion [26], better understanding [27] and efficient team performance [28].

In order for humans to have a seamless, effortless and satisfying interaction with a computer, the process needs to be bidirectional and multichannel. The use of multiple channels in an interaction not only ensures continuity but also adds to the naturalness of this process. We identify verbal, nonverbal, explicit, implicit, intentional and unintentional as the channels used in an interaction process. Use of these channels in human to human interaction has been well researched and documented. However, in HCI especially in the field of HAI, use of multiple channels simultaneously in an interaction needs to be addressed further in order to put HAI on par with human to human interaction.

2.4. Artificial Social Agents

The artificial agents in the form of robots and Virtual Humans (VHs) are becoming more autonomous using the latest technology of artificial intelligence. As robots find their way into our homes with some advantages over animal pets [29], virtual assistants too are becoming more sophisticated and therefore most widely used [1,30]. These artificial agents form the basis for advanced software systems that are being applied in fields such as business, health, and education. Most of these fields require that these agents have the ability to behave in a social context, in addition to assisting and problem solving. These abilities allow agents to become common in everyday environments like homes, offices and hospitals.

As we have seen, for agents to provide seamless, effortless, and satisfactory interaction, it must be capable of multi-channel messages that afford verbal and nonverbal, as well as explicit and implicit communication [31]. This kind of interaction is naturally supported by an artificial social agent, a form of the human-computer interface where the computer communicates with the user through the representation of an animated human form, either as a computer-generated virtual human or a humanoid robot. This type of interaction device will additionally have some degree of agency, allowing them to tailor their feedback behaviour to the specific human they are interacting with. The term Artificial Social Agent (ASA) is used in this paper highlighting the need for these agents to understand, interpret and react using social intelligence as do humans [32].

3. A Bibliometric View

Complementary to the review itself, we conducted a bibliometric analysis of the frequency with which implicit interaction-related keywords occur in the literature. The goal of this analysis is to understand the overall research landscape better. Note that given the narrative nature of the review, the bibliometric dataset did not prescribe what literature was included in the survey itself.

Bibliometric Data

For the bibliometric analysis, the following eight key phrases were searched in Google Scholar: {"human-robot interaction", "'virtual agent' interaction"} × {"proxemics", "gaze", "gestures", "physiology"} i.e., "'virtual agent' interaction proxemics", "human-robot interaction gaze", etc. As a date range, we chose 20 years, from 2000 to 2019. The reason is that the number of publications for the Virtual Agent Interaction (VAI) themes tends to be very low for the years before 2000. In the robotic domain, the term "Human-Robot Interaction" (HRI) is well established, while the terminology used to refer to the interaction with computer-generated, virtual humans is less consistent; terms such as Virtual Agents, Virtual Humans, Virtual Assistants, Embodied Conversational Agents are used interchangeably. For our bibliometric data collection, we only used the term "Virtual Agent Interaction". This limitation might have led to an under-representation of publications in this domain but avoids the issue of including terms that can apply to both domains such as embodied agent and terms that are used for physiological simulations.

Looking at the overall number of publications, we can see that there is a clear difference between the domains of HRI (58,018) and VAI (10,282) (Figure 4a). This five-fold difference between the two domains is quite striking and can only in part be explained by an under-representation of literature in the VAI domain. To accommodate for the overall difference in the number of publications, we normalised the count of publication using the sum of publications per domain (Figure 4b). For both domains, the topics of proxemics and gestures are the least and most mentioned, respectively. The largest relative difference we can find is for the topics of gaze and gestures, where the prior is more prevalent in the VAI domain, and the latter occurred more in the HRI domain.

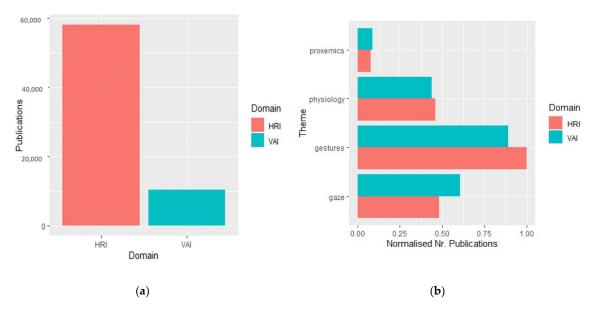


Figure 4. (a) A comparison in the number of publications within HRI and VAI. (b) The number of Google Scholar search results for the use of the terms Gestures, Gaze, Proxemics, and Physiology in the domain of Human-Robot Interaction and Virtual Agent Interaction.

In the span of twenty years, we can see a rapid increase in the number of publications, especially in the HRI domain (Figure 5a). The visualisation using a log scale shows that the ratio of the themes remains surprisingly stable over time (Figure 5b). The exception to this is the increase in publications mentioning proxemics in HRI.

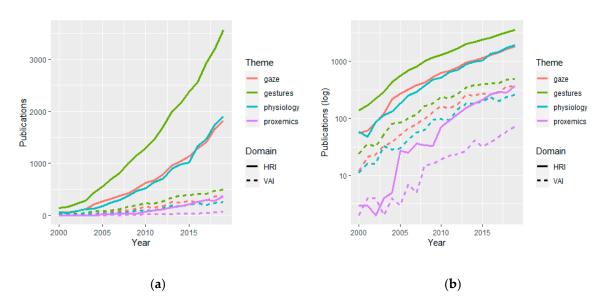


Figure 5. (**a**) In the Human-Robot Interaction domain the number of publications in Gesture, Gaze and Physiology have undergone a rapid increase since 2005 compared to Virtual Agent Interaction. (**b**) The log scale illustrates that this increase has been stable across the modalities.

4. Implicit Interaction between Humans and Artificial Social Agents

To date, a small number of articles have surveyed HRI in nonverbal and a few in implicit domain [33], but to our knowledge, none have looked at ASAs independent of their physical realisation. We argue that humanoid robots and virtual humans do share enough communality in their way of human-like interaction capabilities to treat them as a single category of "Artificial Social Agents". Though currently separated domains, we foresee a confluence between the two types of ASAs as

robots become more physically realistic and virtual reality becomes increasingly immersive. Hence, our review highlights similarities and differences across the two form of physical realisation and identifies areas that have received less attention thus far. In HAI, various modalities of interaction such as gaze, proxemics, gestures, haptics, voice and physiology have been explored previously. We limited our search to interactions using gaze, physiology, proxemics and gestures as they are the most popular in nonverbal domain. The secondary search was carried out using the keywords nonverbal, implicit, and interaction. Our search focused on the most current academic journals and conference papers published in the English language and primarily within computer science. The subsequent review highlights the work on HRI and VAI within the most common modalities of Gesture, Gaze, Proxemics, and Physiology.

4.1. Physical Distance and Body Orientation

Proxemics is the interpersonal distancing behaviour that people observe when interacting with each other. Hall divides the space around an individual into several layers ranging from intimate space—which is the nearest, to public space—which is the farthest from an individual [34]. The proxemic hypothesis postulates that the placement of an individual within one's intimate space evokes feelings of invasion and discomfort while as the placement of an individual within someone's personal space increases persuasiveness. Therefore, proxemics plays an important role in determining how an individual is perceived by others.

Research on interaction with an ASA suggests that incorporating nonverbal cues like proxemics improve the persuasiveness and peoples compliance with the robot [35]. Moreover, robots that do not observe appropriate distancing behaviour are perceived as threatening and disruptive to social environments and work practices [36]. This could be because the robot's proxemic behaviour affects peoples emotional state and their perception of the robots social presence [37]. These results correspond to research in human to human interaction and suggest that robots interpersonal distancing behaviour determines how they are perceived by people. Proxemics research, in terms of avoidance rules in a path crossing, reports that robots that observe normal path crossing rules are treated similar to humans [38] and instil positive emotional response in people [37]. Similarly, people expressed annoyance and stress when the robot regarded them as a static obstacle [39].

In another study, the likeability and gaze of the robots also affected the users' proxemic behaviour. Robot's gaze at the users' leads to an increase in distance between them and the robot if they dislike the robot [40]. Likeability furthermore is associated with information disclosure. People disclosed less personal information to the disliked robot and more to the robot they liked [41].

Body orientation has been studied in ASA interaction. When humans communicate, they orient their gaze towards the focus object. However, if they continue observing the focus object, they align their body orientation towards that focus object, creating a space where their attention focuses together. Kendon, calls this "O-Space" [42] and suggests that body position and orientation are the important indicators of continuous attention while gaze and pointing behaviours indicate instantaneous attention [43]. Yamaoka et al. developed a model based on the concept of O-Space, where they suggest that people tend to stand close to their partner and the focus object [44,45]. This, however, was found to be challenging when a robot was presenting information, where, considering robots own as well as the listeners' field of view was important. Therefore, the authors suggest the following constraints when interacting; proximity to listener, proximity to the object, listener's field of view, and presenter's field of view [46].

The proxemic behaviour in human and virtual human interaction reveals that people observe similar distancing behaviour with a VH as they do with a real human. When people interact, they orient their posture according to their interaction partner. Similar observations were recorded in human and virtual human interaction where users' changed their posture and orientation in line with their interaction partner [47]. The presence of a VH can lead to difficulty in remembering information, indicating the presence of the phenomenon of social facilitation [48]. Multimodal studies

using proxemic and gaze are studied in human and virtual human interaction too. The direct gaze of a VH led to an increase in spacing between the VH and the user [48]. Users' left more distance in between when approaching the VH from the front compared to the rear of the VH. This study further reports that gender is associated with proxemic behaviour. Female participants left more interpersonal distance between them and the VH as compared to men [49].

In sum, research on proxemic behaviour in human agent interaction suggests that if agents follow the human to human proxemic behaviours as described by Hall [34], people not only perceive them as socially more present but also as more positive compared to the agents that do not adhere to the human to human proxemic rules. For example, avoiding getting too close to people by entering their personal and intimate space and giving "right of way" in a path crossing scenario. By adhering to human–human proxemic rules, agents are more persuasive, and people comply with them better compared to an agent that treat people as static objects.

The nature of the task given to participants modules proxemic behaviour; participants playing a card game with a robot in two scenarios: competing with the robot and cooperating with the robot. The participants' found a positive experience when competing with an agent when it was close compared to when they are cooperating with a distant robot [50].

Another observation is that peoples' characteristics influence their interpersonal distance with an agent. People who scored low on agreeableness leave more space compared to those who scored high. Female participants tend to leave more distance between them and the agent. Pet owners are comfortable moving closer to the agent and letting the agent come closer to them.

Looking at it from the opposite direction, properties attributed to the agent such as likeability also affect proxemics: Participants leave more interpersonal distance from a disliked agent. With the "Uncanny Valley" hypothesis, Mori [51] predicts a specific, non-linear relationship between the likeability and the likeness. The uncanny valley is the phenomenon where the human likeness of an agent provokes feelings of dislike in people after a certain threshold. It is a dip in the positive relationship between an agent's likeness and likeability.

Based on the above, we can say that the likeness of an agent guides likeability which in turn guides proxemics. However, studies show that likeability can change based on peoples expectation [52], and people's expectations change based on likeness [53], such that as the degree of anthropomorphism increases, so does the people's expectations [54]. This suggests that the relationship between likeability and proxemics is mediated by likeness and user expectations. Given this complexity, the relationship between user expectations, likeness, likeability and proxemics needs further exploration.

4.2. Interaction through Gestures and Facial Expression

Human to human interaction relies heavily on nonverbal cues like facial expressions and body movements [55,56]. Gestures are a type of nonverbal movements used to convey an explicit message or accompany a verbal message. Gestures not only express a command or an instruction but a specific execution of the gesture can be an indicator of the person's psychophysiological states [57]. Although the use of nonverbal cues such as gestures make the interaction simpler and natural [58], culturally diverse gesture styles and the variety of words and pointing gestures used to indicate the same objects makes gesture recognition a challenge in ASA interaction.

Gesticulation was incorporated in BERTI (Bristol and Elumotion Robotic Torso 1) where it could exhibit three gestures: a follow-me gesture, where the robot raises its arm, palm upward, towards its face, makes two strokes, then retracts its arm down; a take gesture, where the robot raises a closed fist to the centre of its torso, then extends its hand, then retracts its arm down and; a shake-hand gesture where the robot moves its right arm, palm facing to the left up, and then fully forward, then retracts its arm down [59]. These gestures are based on findings that suggest human gestures consist of three phases: preparation, stroke, and retraction [43]. Multi-modal gesticulation systems have been reported to make human-robot interaction simple and natural [58].

Another important bodily mechanism used to display emotions is facial expressions [60]. Incorporating naturalistic facial expressions can be challenging as robot's do not have muscles like humans and there is a risk of falling into the uncanny valley [61]. In an attempt to tackle this, Kobayashi and Hara created a robot with facial expressions for surprise, fear, disgust, anger, happiness, and sadness [62]. Later on, robot Kismet was made that was also capable of producing facial expressions and changes in body posture [63]. Some robots used a screen to display their facial gestures with an animated face on the screen like Robovie-X [64], Olivia [65] and Iromec [66].

While in human-robot interaction, use of gestures increases the robot's persuasiveness [35], it created a sense of rapport between the user and the agent in human virtual–human interaction [67]. Another similar animation system is available in human and virtual human interaction, which can produce automated non-verbal behaviours such as nodding, gesturing and eye saccades [68]. Use of verbal as well as nonverbal cues of gestures and gaze in a multimodal study has also been reported for turn-taking [69]. However, studies in human virtual–human interaction that solely investigates the role of gestures are sparse, even though these developments can enhance the expressiveness and aid communication between a human and virtual human [70].

In summary, ASAs capable of realistic gesticulation and facial expression afford an easier and faster interaction. Moreover, adding other communication channels such as gaze has been found to increase the realism of ASA interaction. While most of the studies solely rely on the verbal report and subjective ratings by the participants, Breazeal's study reports elicitation of unconscious mirroring gesticulation behaviours between humans and the robot, where the user and the robot unconsciously followed each other's head gestures iteratively [63].

Most of the studies investigating gesticulation are limited in scope as they only consider a subset of the four gesture types, focusing either on dietic [71], beat [72], iconic [73], or metaphoric gestures [74]. A notable exception is Chidambaram and colleagues who looked at all four gesture types [35].

The development in Human–Robot Interaction in terms of realism of facial expressions has come a long way from Iromec, presented in 2010 [66], to Sophia, presented in 2019 [75]. However, state-of-the-art robotic faces still evoke the negative responses towards these human-like realistic-looking robots [61,76] indicating the presence of the uncanny valley. Interestingly, we can witness here a confluence between the humanoid robotic and virtual humans domains; higher levels of realism in Virtual Humans equally has shown the presence of the uncanny valley effect [77,78].

4.3. Communicative Gaze Behaviour

Gaze is the act of looking at something. It is a natural form of interaction where the eye movements are steady and intentional. It is one of the channels used in an interaction that conveys implicit messages such as attention, liking and turn-taking [79,80].

There are several facets of gaze discussed in the research, including mutual gaze, referential gaze, joint attention, and gaze aversions [81]. Mutual gaze is the social phenomenon where both the interacting entities are gazing at each other [82]. However, people often use gaze aversions when interacting with one another. This is when the gaze is shifted from the eyes to other parts of the partners face. In Referential gaze, the gaze is averted towards a common reference or object. Joint attention is another facet of gaze where interaction partners start with a mutual gaze, followed by a referential gaze to draw attention towards a common object, and then back to a mutual gaze in some cases. Establishing an eye gaze sends implicit cues for the opening of the communication channel between people [83] where they use their observations of others' eye gaze to obtain implicit cues on the willingness to carry on or terminate the conversation.

Das and Hasan suggest that gaze is one of the most important aspects of an agent as it is found to elicit a similar response in the interaction between humans and ASAs as is in human to human interaction [84]. For example, people exaggerate their actions to re-engage the robot's gaze, implying robots gaze conveys implicit cues on attention [85]. Another interesting finding was reported where users' in a storytelling session remembered the plot better when the robot used its gaze compared

to when the robot did not look at them [86]. They further found differences in male and female where men evaluated robot more positively.

Gaze is studied widely along with proxemics in ASA interaction. Users' maintain a greater distance with the robot that maintains a mutual gaze with them [40]. In such multimodal studies, robot's gaze behaviour was found to aid in persuasiveness [35] and facilitate user's understanding of the robotic speech [87], leading to better interaction with the robot [88]. Another interesting observation is that pet owners, who are adept at nonverbal communication with their animals, are more sensitive to nonverbal behaviours, and therefore can exploit information from robotic gaze better than the people who are not pet owners [88]. Additionally, they took significantly less time and required fewer questions compared to people who are not pet owners to complete a task with a robot, further indicating that training could lead to an effective inference of nonverbal cues.

As eye movement requires minimal physical effort by the user, gaze behaviour has been widely researched in the context of elderly care and assistive technology [89,90]. Li and Zhang designed a framework to provide assistive systems with the capability of understanding the user's intention implicitly by monitoring their overt visual attention [91]. For example, if the user gazes at a specific location for a pre-determined amount of time, the robot interprets it as an implicit signal conveying the need for further information [92]. With the intention of monitoring implicit cues of the user's attention, Lahiri and colleagues [93] present a similar system where the user and the agent are in a dynamic closed-loop interaction based on real-time eye tracking. Depending on the level of user engagement via their gaze behaviour, the agent changes its mode of talking to increase interactivity. Another study based on the similar closed-loop interaction is by Bee and colleagues, where the agent's gaze is changed in response to users gaze [94]. Next to these closed-loop interactions, there are studies that present feed-forward open-loop interactions, e.g., [95]. Here, the category of the agent's gaze behaviour is pre-defined. The categories tested are gaze only, gaze with eye blinks and head nods, and intermittent gaze at the user. The study reports that the condition where the agent gazed and performed the actions such as eye blinks and head nods convey interest from the agent while the rest of the conditions convey disinterest. The study further reports that disinterest from agent leads to a decreased rapport between the user and the agent. These findings are in line with another study where they report that if the agents gaze is responsive based on the user's gaze, it increases the users' sense of social presence and rapport [96].

Set in human-robot interactions, a similar feedback approach has been implemented where based on user's proxemic behaviour, the robot changes its gaze behaviour. The robot's "idle state" condition is triggered when the user is not in the conversational area of the robot. In this state the robot fixes its gaze to a pre-defined location, stays idle and waits for the user to come closer to it; The "listening state" is triggered when the user moves into the conversational area of the robot. Here the robot orients its body towards the speaker and maintains eye contact while listening to the user; In the "speaking state"—the robot orients its body, maintaining eye contact while speaking to the user and; In "pointing state"—the robot points out to the reference object and changes its gaze towards that object [97].

Similar to human–robot interactions, multi-modal studies combining gaze behaviour and proxemics have been reported in human virtual-human interactions. Changes in gaze led to changes in the physical distances that people keep between themselves and the VH [48]. An interesting finding was the effect of gender on distancing behaviour: Female participants were found to increase the distance between themselves and the VH if the agent maintained a constant gaze. These results are in line with human–robot interactions, where robots gaze oriented towards the human increased the distance humans maintained between themselves and the robot [40]. Therefore, gaze and proxemic behaviour is fluence each other. However, Fiore and colleagues report that proxemic behaviour is more influential than the gaze behaviour of the robot [37]. Corroborating this finding, Mumm and Mutlu report that adding gaze has a contrasting relationship with proxemics—meaning that users increase the distance between them and the agent if the agent fixed its gaze at the user and vice versa.

They also bring in the likeability variable reporting that if the users disliked the robot, the gaze from the robot made them increase the distance to the robot [40].

Implementing realistic eye behaviour is a challenging task as this not only consists of the location of gaze fixations, their duration, and the speed of saccades but also involves subtle effects such as pupil's reaction to light pupil dilation and squinting. Furthermore, the anatomical intricacies in the eye such as the layered structure of the iris add to the complexity of modelling the visual appearance of eyes.

Thus far, we have discussed the interaction between humans and ASAs based on salient bodily cues such as distance and gestures. What about the subtle, implicit interaction cues such as blushing and breathing? The next section focuses on interaction based on such "physiological" responses.

4.4. Interaction Based on Physiological Behaviour

The implicit cues that people send and receive in an interaction are often unintentional and very subtle—to the point that neither the receiver nor the sender is aware of these cues. However, their physiological responses via GSR, heart rate, skin conductance measurements can give away the effect of the implicit influence. Research in human to human interaction suggests that romantic partners send physiological cues that lead to heart rate adaptation in partners [16,17]. In HCI, we can use a similar physiological response such as breathing rate using breathing belts or thermocouple to record the effect of influence between humans and computers. In human to human interaction, the phenomenon of interacting people unintentionally and unconsciously mimicking each other is the Chameleon Effect [19]. The physiological influence causes an iterative effect on the interaction partners, leading to a state of synchrony between them [15]. This physiological synchrony is pivotal in nonverbal human–human interaction that helps building and maintaining rapport between people [20]. Therefore, in HCI we are seeing a trend in building systems that mimic the synchronisation process that exists in human to human interaction. For example, a study reports mirroring between human and robot [63]. As the robot lowers its head in response to the user's instruction, the user responds by lowering their head. It was observed that the cycle continues to increase in intensity until it bottoms out with both subject and robot having dramatic body postures and facial expressions that mirror each other. The humanoid robot Nico synchronizes its drumming to that of another person or a conductor [98]. Nico uses visual and auditory channels to gauge the conductor's movements to match their physiology. In a similar study, a robot accompanist can start and stop performance in synchrony with a flutist, and can detect tempo changes in real time [99].

Biofeedback has been used as an implicit cue, where a robot assesses the anxiety levels of human participants during a baseball game and acts as a coach in the game. Depending on the anxiety level of the participant, the robot changes the difficulty level of the game [100,101] Participant's physiological signals were measured through wearable biofeedback sensors and analysed for affective cues which were inferred by a robot in real time. The study reports an improved performance based on anxiety reduction in participants.

Another study that measured physiological data was by Mower et al., where the users played a wire puzzle and the game was moderated by either a simulated or embodied robot [102]. The data were used by the robot to monitor the users' progress and level of disengagement. Based on the physiological data of GSR and skin temperature, the level of the game was appropriately adapted to re-engage the user. In virtual reality-based interactions, physiological responses are reported to be related to the psychological state where changes in head movements correspond to social anxiety in a virtual classroom [103]. Furthermore, such responses from users during a relaxation session were evaluated based on the virtual instructor compared to a video-based human instructor [104]. The results revealed positive responses from the users reporting lower anxiety levels and higher mindfulness levels with the virtual instructor compared to a human instructor. The study was modified further where the virtual instructor was aware of the users breathing and gave regular feedback to the users [105] to improve the stress reduction and mindfulness. Moreover, the phenomenon of social facilitation—where the task

performance varies depending on the presence or absence of another person—has been reported in human virtual–human interaction [106].

A fair amount of research has been done on nonverbal agent behaviours, and some have explored explicit VAI [104], some implicit behaviours in HRI [107], while some suggest that combining both explicit and implicit cues in an interaction will lead to a better HAI [108].

The systems we have looked at thus far record physiological activity as an implicit signal that is interpreted by the computer to assess an internal state of a user. These systems then either use this signal is to adapt the agent's behaviour or store it for off-line analysis of the quality of the interaction. An alternative and rarely explored application is the use of uses of agent's implicit "physiological" cues to influence the same behaviour in the user to induce, what in the psychological literature is referred to as, physiological synchrony.

In the next section, we will illustrate the use of this configuration by presenting a case study of our novel framework in which a virtual human actively entrains the physiological behaviour, namely breathing, of a user.

5. Case Study: The Virtual Human Breathing Relaxation System

We present a case study on our system that harnesses a bi-directional implicit interaction between a human and a VH. Our system explores the implicit behaviours through physiological synchrony and builds on work from different areas: bio-responsive systems, immersive virtual reality and breathing entrainment.

We focus on breathing as it has a bidirectional relationship with emotion and cognition [109,110] and hence improving breathing has a positive effect on mood and improve mental health [111,112]. Several breathing relaxation training exercises are being prescribed to patients with mild to severe breathing condition like asthma, stress, anxiety, insomnia, panic disorder, chronic obstructive pulmonary disease, etc. [113,114]. Virtual Reality (VR) has been used in treating some of the disorders mentioned above such as anxiety, depression, schizophrenia, phobias, post-traumatic stress, substance misuse and eating disorders [115,116]. VR has also been found useful in relaxation therapies [104]. However, the systems developed so far heavily rely on explicit, verbal instruction given to the user: In a typical VR breathing system, a virtual therapist guides the patients breathing into synchrony with their breathing eventually leading the patient into a relaxed state. Contrary to these approaches, our system harnesses the established phenomenon in human to human interaction physiological synchronisation. Physiological synchrony also referred to as entrainment is where the oscillation of one system is being influenced by the oscillation of another.

5.1. Physiological Synchrony Based on Active Entrainment

Our system is an automated breathing relaxation system that harnesses the known mechanisms of physiological synchrony through active entrainment. In active entrainment, the oscillations of one system are influenced by the oscillations of the other system. Both systems influence each other iteratively in a closed loop and which results in a synchronised oscillation (Figure 6). Hence, in a system with active entrainment, the oscillations of the target system are recorded, and the controller system then adapts the rhythm of the external pacemaker such that it can change the target rhythm gradually over time, by continuously delivering a signal of optimal influence.

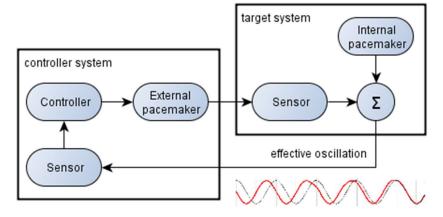


Figure 6. The concept of active entrainment is at the core of the Virtual Human Breathing Relaxation System (VHBRS). The breathing of a Virtual Human serves as the external pacemaker that influences the breathing rhythm of the target system, i.e., the user.

In VHBRS, a finite state machine is used to implement the active entrainment: Control system starts at a pre-set breathing rate, and the target system (i.e., the user) breathes at a different breathing rate as illustrated in Figure 7. Based on the target system's breathing frequency, the control system, in turn, adapts the breathing frequency. Once synchrony is achieved, the control system changes the breathing frequency gradually, "dragging along" the target systems breathing rhythm until the desired frequency is achieved. Note that the proposed concept of active entrainment is generic and therefore can be applied to other physiological measures such as GSR or ECG.

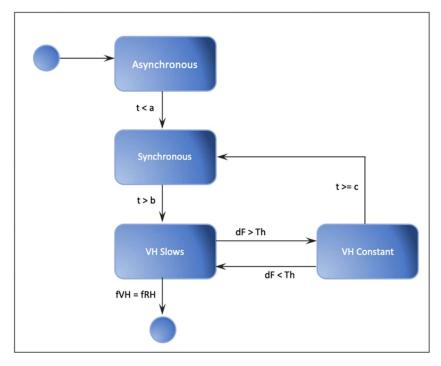


Figure 7. Active entrainment using a finite state machine. Initially, user and VH are breathing at different rates (Asynchronous). After time a, the VH changes its breathing frequency to match the participants (Synchronous). After time b, the VH slows down its breathing frequency. Once the threshold difference Th between the two breathing frequencies is reached, the VH breaths at a constant frequency. Depending on whether the threshold is reached, the VH either re-synchronises with the user or remains at its current breathing frequency.

System Implementation

On the input side, we are recoding breathing behaviour of the participants using two belts connected to a BITalino (https://bitalino.com) device. The immersive virtual environment and the virtual human in VHBRS are created and rendered using the game engine Unity (https://unity3d.com) as shown in Figure 8. The breathing sounds are controlled using the audio synthesis platform SuperCollider (https://supercollider.github.io/). The real-time data processing and finite-state machine are implemented in the Matlab-based tools Simulink (https://uk.mathworks.com/products/stateflow.html), respectively. The middleware ActiveMQ (http://activemq.apache.org/) is used for communication between the components.

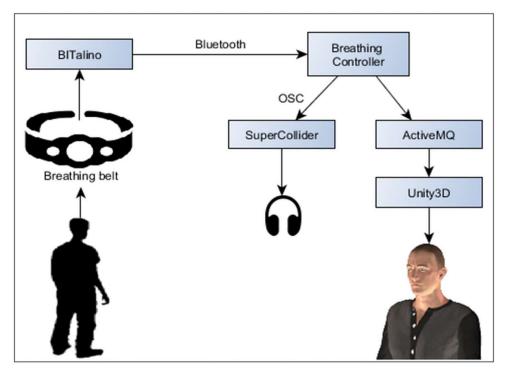


Figure 8. Architecture of the Virtual Human Breathing Relaxation System (VHBRS).

5.2. Pilot Study

The full-fledged VHBRS is currently under development. In the first pilot study, we aimed to determine the feasibility of using breathing signals from a VH, namely body movement and breathing sounds, to influence the breathing behaviour of participants [117]. As a step toward the closed-loop control system, we implemented an open-loop system that uses passive entrainment, meaning that the external pacemaker is set to a fixed frequency independent of the frequency of the target system (Figure 8). In the pilot study, participants were instrumented with two sensor belts to record chest and abdominal breathing and placed in an immersive virtual environment where they encountered a VH (Figure 9). Before the exposure, participants were not given any specific instruction as to what to do and were not informed about the specific purpose of the study.

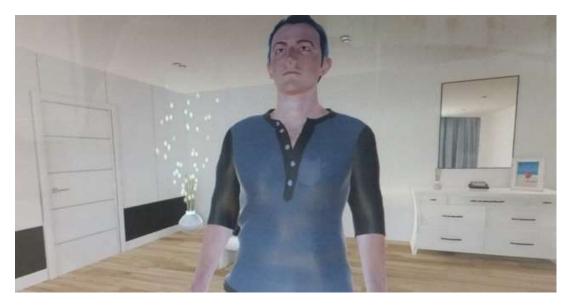


Figure 9. The immersive virtual reality environment as seen by the participant with the VH standing in a living room.

One of the issues with open-loop systems is that the breathing frequency of the VH and the participant may naturally coincide. To avoid this caveat, the breathing frequency of the VH was changed repeatedly during a single trial. Specifically, the VH's breathing frequency was alternated between two frequencies six times as shown in Figure 10. Each of the phases was two minutes long.

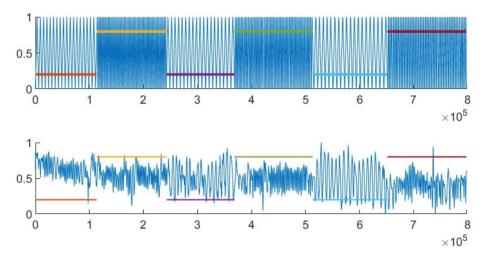


Figure 10. Top plot: Breathing frequency of the virtual human during a single pilot trial. Bottom plot: Example breathing recording from a human participant.

Findings

Twelve university students participated in the pilot study (10 males and 2 females). The participants were contacted via emails. To assess the influence of the VH breathing behaviour on the participants, we calculated the participants' breathing frequency for each of the six phases of the trial (Figure 11). A visual inspection of the breathing shows that some participants indeed showed the expected acceleration and deceleration of the breathing frequency. The occasionally spurious data are due to issues with the position of the sensor belts and show the challenges associated with recording breathing data reliably using this method.

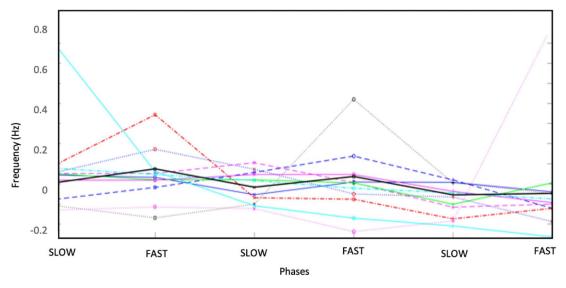


Figure 11. Abdominal breathing frequency of all participants during the six phases of the trial.

To get feedback on the environment and the VR experience, a post-experimental questionnaire was used. Three items assessed prior VR experience or motion sickness, while five items pertained to implicit interaction experience. Results for the latter are shown in Figure 12. The data show that the participants reported feeling more relaxed after the interaction with the VH, and a majority said that they noticed the VH's breathing. All participants reported noticing the change of breathing rate of the VH. Interestingly, only half of the participants reported being influenced explicitly by the VH while five participants reported that they were actively trying to synchronise with the VH, even though no explicit instructions about breathing were given to the participants.

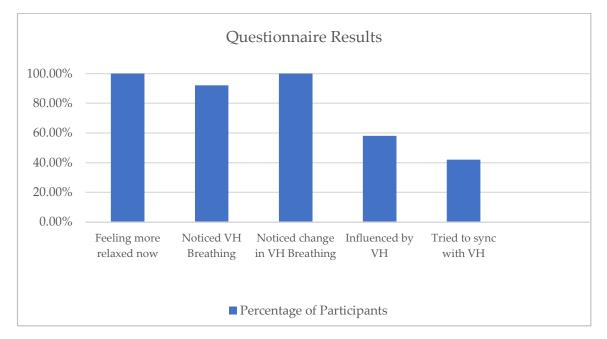


Figure 12. Post-experimentation questionnaire results.

5.3. Case Study Conclusion

From a technical point of view, we have learned that breathing belts are not robust against changes in posture. A thermistor-based system that is worn under the nostrils, would provide more robust data, though this is can make data collection more difficult if there are concerns of the spread of pathogens. The questionnaire data shows that the participants were picking up the implicit signal, likely also because instrumentation with breathing belts makes it clear what cues to look out for. If this is not desirable and only an implicit, unintentional influence is wanted, one would need to recur to some level of deception, e.g., by using additional sensors such as GSR and ECG to make it less clear that breathing is the focus.

Users accepted the expression of implicit "physiological" signals—breathing movement and sounds—from a computer-generated human, though sensibly, the agent does not require the physiological processes. It would, in this regard, be interesting to see how users react to a more "corporeal" entity, i.e., a robot.

Interestingly, even though no stress was explicitly induced at the beginning of the experiments, participants report in increase is relaxation. In a continuation, this effect could be amplified by first submitting participants to a stress-inducing task, to then investigate if the VHBRS is able to accelerate the return to pre-stress baseline. This stress induction would, however, require careful ethical consideration to make sure participants indeed do not leave the experiment more anxious than beforehand.

Taken together, we are confident that the full VHBRS which will include real-time, active entrainment, will provide an effective framework for influencing the user's levels of relaxation.

6. Conclusions

In this paper, we identified seamlessness, effort, and satisfaction as the key factors of natural human-computer interaction. We took reference from human to human interaction and from human to human-mediated interaction to highlight that interaction is a bi-directional and multichannel process where the actions of one influence the actions of others. The result is a reciprocal action that the receiver takes based on the effect of sender's cues. These cues are transmitted over multiple channels that can be classified along the dimensions of verbal–nonverbal, explicit–implicit, and intentional–unintentional.

The use of multiple channels in an interaction process is so effortless that people often do not realise that they are using all these channels while interacting. The use of the implicit channel is often unintentional further reducing the effort required to interact, thereby making the process more efficient. This is supported by research that suggests incorporation of implicit cues makes the interaction unambiguous [25] and efficient [26]. These unintentional and implicit cues are frequently the drivers of physiological responses and therefore pivotal in an interaction. In human to human interaction, implicit cues lead to change in physiological state with people imitating their partner's physiology leading to an increase in rapport and bonding between them.

Although HAI has come a long way, it is still not on par with human to human interaction in terms of feeling completely natural and efficient. Therefore, we reviewed the literature to highlight the progress of research on incorporating the multiplicity of channels and on the feedback for reciprocal action. The literature was reviewed on nonverbal channels with special focus on implicit interactions between a human and artificial social agent in the form of robots and human and virtual humans.

We focused on the main modalities that capture the implicit interactions namely—Gestures, Proxemics, Gaze and Physiology. Our bibliometric analysis shows that overall, Human-Robot Interaction (HRI) is more popular than Virtual Human Interaction (VAI). We can see a trend that the modality of Gestures is the most popular in both domains. The literature review shows that incorporating gestures in the interaction between humans and artificial social agents aids the interaction process by increasing the agent's persuasiveness and virtual humans' expressiveness. Proxemic behaviour of agents can instil the same emotional response in the user as do other humans. Although proxemics seems to be more relevant to robots, VH too also elicits responses like a human, especially when mixed with another modality such as Gaze.

Surprisingly, research in all the modalities reports that humans treat ASA and other humans the same with similar expectations from both. People show an emotional response to these agents similarly as they do to other humans, e.g., getting offended if they are not given their right of way

while crossing a road or feel intimidated if they are being constantly gazed at. Humans also show compliance to these agents if they demonstrate the "humanly" nonverbal cues while interacting and highlight that agents have the potential to persuade people if they use appropriate nonverbal cues. The consequence of these behaviours is a positive attribution to the perception of ASA's if they comply with the rules that apply in a human to human interaction.

Considering the importance of nonverbal and in specific implicit information exchange in an interaction, it is no surprise that more research needs to be done within the implicit interaction domain of HAI. The current literature review accentuates that not many computer systems have been developed with implicit interaction capability.

In the last section of this paper, we present an example of how such a system would function using nonverbal, implicit, and unintentional channels within HAI. Our system showcases mechanisms that are found in human to human interaction. Our system is called the Virtual Human Breathing Relaxation System and can harness a more naturistic interaction through physiological synchrony and active entrainment between a human and a virtual human.

Author Contributions: S.D.: Initial draft of literature review, Design, Implementation and Execution of the pilot study; U.B.: Conceptualisation of the pilot study, Writing and Editing literature review. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

References

- Griffin, J. Voice Statistics for 2020. 2020. Available online: https://www.thesearchreview.com/google-voice-searches-doubled-past-year-17111/ (accessed on 9 July 2020).
- 2. IDTechEx. Smart Speech/Voice-Based Technology Market Will Reach \$ 15.5 Billion by 2029 Forecasts IDTechEx Research. 2019. Available online: https://www.prnewswire.com/news-releases/smart-speechvoice-based-technology-market-will-reach--15-5-billion-by-2029-forecasts-idtechex-research-300778619.html (accessed on 20 April 2020).
- Kiseleva, J.; Crook, A.C.; Williams, K.; Zitouni, I.; Awadallah, A.H.; Anastasakos, T. Predicting user satisfaction with intelligent assistants. In Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval, Pisa, Italy, 17–21 July 2016; pp. 45–54.
- 4. Cox, T. Siri and Alexa Fails: Frustrations with Voice Search. 2020. Available online: https://themanifest.com/ digital-marketing/resources/siri-alexa-fails-frustrations-with-voice-search (accessed on 16 May 2020).
- 5. Major, L.; Harriott, C. Autonomous Agents in the Wild: Human Interaction Challenges. In *Robotics Research*; Springer: Cham, Switzerland, 2020; Volume 10, pp. 67–74.
- Ochs, M.; Libermann, N.; Boidin, A.; Chaminade, T. Do you speak to a human or a virtual agent? automatic analysis of user's social cues during mediated communication. In Proceedings of the ICMI—19th ACM International Conference on Multimodal Interaction, Glasgow, UK, 13–17 November 2017; pp. 197–205.
- Glowatz, M.; Malone, D.; Fleming, I. Information systems Implementation delays and inactivity gaps: The end user perspectives. In Proceedings of the 16th International Conference on Information Integration and Web-Based Applications & Services, Hanoi, Vietnam, 4–6 December 2014; pp. 346–355.
- 8. Schoenenberg, K.; Raake, A.; Koeppe, J. Why are you so slow?—Misattribution of transmission delay to attributes of the conversation partner at the far-end. *Int. J. Hum. Comput. Stud.* **2014**, 72, 477–487. [CrossRef]
- 9. Jiang, M. The reason Zoom calls drain your energy, BBC. 2020. Available online: https://www.bbc.com/ worklife/article/20200421-why-zoom-video-chats-are-so-exhausting (accessed on 27 April 2020).
- Precone, V.; Paolacci, S.; Beccari, T.; Dalla Ragione, L.; Stuppia, L.; Baglivo, M.; Guerri, G.; Manara, E.; Tonini, G.; Herbst, K.L.; et al. Pheromone receptors and their putative ligands: Possible role in humans. *Eur. Rev. Med. Pharmacol. Sci.* 2020, 24, 2140–2150. [PubMed]
- 11. Melinda, G.L. Negotiation: The Opposing Sides of Verbal and Nonverbal Communication. J. Collect. Negot. Public Sect. 2000, 29, 297–306. [CrossRef]
- 12. Abbott, R. Implicit and Explicit Communication. Available online: https://www.streetdirectory.com/etoday/ implicit-andexplicit-communication-ucwjff.html (accessed on 20 July 2020).

- 13. Implicit and Explicit Rules of Communication: Definitions & Examples. 2014. Available online: https://study. com/academy/lesson/implicit-and-explicit-rules-of-communication-definitions-examples.html (accessed on 3 July 2020).
- 14. Thomson, M.; Murphy, K.; Lukeman, R. Groups clapping in unison undergo size-dependent error-induced frequency increase. *Sci. Rep.* **2018**, *8*, 1–9. [CrossRef] [PubMed]
- Palumbo, R.V.; Marraccini, M.E.; Weyandt, L.L.; Wilder-Smith, O.; McGee, H.A.; Liu, S.; Goodwin, M.S. Interpersonal Autonomic Physiology: A Systematic Review of the Literature. *Personal. Soc. Psychol. Rev.* 2017, 21, 99–141. [CrossRef] [PubMed]
- 16. McAssey, M.P.; Helm, J.; Hsieh, F.; Sbarra, D.A.; Ferrer, E. Methodological advances for detecting physiological synchrony during dyadic interactions. *Methodology* **2013**, *9*, 41–53. [CrossRef]
- 17. Ferrer, E.; Helm, J.L. Dynamical systems modeling of physiological coregulation in dyadic interactions. *Int. J. Psychophysiol.* **2013**, *88*, 296–308. [CrossRef]
- 18. Feldman, R.; Magori-Cohen, R.; Galili, G.; Singer, M.; Louzoun, Y. Mother and infant coordinate heart rhythms through episodes of interaction synchrony. *Infant Behav. Dev.* **2011**, *34*, 569–577. [CrossRef]
- 19. Chartrand, T.L.; Bargh, J.A. The chameleon effect: The perception-behavior link and social interaction. *J. Pers. Soc. Psychol.* **1999**, *76*, 893–910. [CrossRef]
- 20. Lakin, J.L.; Jefferis, V.E.; Cheng, C.M.; Chartrand, T.L. The chameleon effect as social glue: Evidence for the evolutionary significance of nonconscious mimicry. *J. Nonverbal Behav.* **2003**, *27*, 145–162. [CrossRef]
- 21. Severin, W. Another look at cue summation. A.V. Commun. Rev. 1967, 15, 233–245. [CrossRef]
- 22. Moore, D.; Burton, J.; Myers, R. Multiple-channel communication: The theoretical and research foundations of multimedia. In *Handbook of Research for Educational Communications and Technology*, 2nd ed.; Jonassen, D.H., Ed.; Lawrence Erlbaum Associates Publishers: Mahwah, NJ, USA, 1996; pp. 851–875.
- 23. Baggett, P.; Ehrenfeucht, A. Encoding and retaining information in the visuals and verbals of an educational movie. *Educ. Commun. Technol. J.* **1983**, *31*, 23–32. [CrossRef]
- 24. Mayer, R.E. Multimedia learning. Psychol. Learn. Motiv. Adv. Res. Theory 2002, 41, 85–139.
- 25. Adams, J.A.; Rani, P.; Sarkar, N. Mixed-Initiative Interaction and Robotic Systems. In *AAAI Workshop* on Supervisory Control of Learning and Adaptative Systems; Semantic Scholar: Seattle, WA, USA, 2004; Volume WS-04-10, pp. 6–13.
- 26. Blickensderfer, E.L.; Reynolds, R.; Salas, E.; Cannon-Bowers, J.A. Shared expectations and implicit coordination in tennis doubles teams. *J. Appl. Sport Psychol.* **2010**, *22*, 486–499. [CrossRef]
- Breazeal, C.; Kidd, C.D.; Thomaz, A.L.; Hoffman, G.; Berlin, M. Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems IROS, Edmonton, Canada, 2–6 August 2005; pp. 708–713.
- Greenstein, J.; Revesman, M. Two Simulation Studies Investigating Means of Human-Computer Communication for Dynamic Task Allocation. *IEEE Trans. Syst. Man. Cybern.* 1986, 16, 726–730. [CrossRef]
- 29. Coghlan, S.; Waycott, J.; Neves, B.B.; Vetere, F. Using robot pets instead of companion animals for older people: A case of "reinventing the wheel"? In Proceedings of the 30th Australian Conference on Computer-Human Interaction, Melbourne, Australia, 4–7 December 2018; pp. 172–183.
- 30. Google Inc. Teens Use Voice Search Most, Even in Bathroom, Google's Mobile Voice Study Finds. 2014. Available online: https://www.prnewswire.com/news-releases/teens-use-voice-search-most-even-in-bathroom-googles-mobile-voice-study-finds-279106351.html (accessed on 21 April 2020).
- 31. Kang, D.; Kim, M.G.; Kwak, S.S. The effects of the robot's information delivery types on users' perception toward the robot. In Proceedings of the 2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), Lisbon, Portugal, 28 August–1 September 2017; pp. 1267–1272.
- 32. Schmidt, A. Implicit human computer interaction through context. *Pers. Ubiquitous Comput.* **2000**, *4*, 191–199. [CrossRef]
- 33. Saunderson, S.; Nejat, G. *How Robots Influence Humans: A Survey of Nonverbal Communication in Social Human–Robot Interaction;* Springer: Dordrecht, The Netherlands, 2019; Volume 11.
- 34. Hall, E. The Hidden Dimension; Doubleday: Garden City, NY, USA, 1966.
- 35. Chidambaram, V.; Chiang, Y.; Mutlu, B. Designing persuasive robots: How robots might persuade people using vocal and nonverbal cues. In Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction, Boston, MA, USA, 5–8 March 2012; pp. 293–300.

- 36. Mutlu, B.; Forlizzi, J. Robots in organizations. In Proceedings of the 2008 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI), Amsterdam, The Netherlands, 12–15 March 2008; p. 287.
- Fiore, S.M.; Wiltshire, T.J.; Lobato, E.J.C.; Jentsch, F.G.; Huang, W.H.; Axelrod, B. Toward understanding social cues and signals in human-robot interaction: Effects of robot gaze and proxemic behavior. *Front. Psychol.* 2013, *4*, 1–15. [CrossRef]
- 38. Vassallo, C.; Olivier, A.H.; Souères, P.; Crétual, A.; Stasse, O.; Pettré, J. How do walkers behave when crossing the way of a mobile robot that replicates human interaction rules? *Gait Posture* **2018**, *60*, 188–193. [CrossRef]
- Dondrup, C.; Lichtenthäler, C.; Hanheide, M. Hesitation signals in human-robot head-on encounters: A pilot study. In Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot, Interaction, Bielefeld, Germany, 3–6 March 2014; pp. 154–155.
- Mumm, J.; Mutlu, B. Human-robot proxemics: Physical and Psychological Distancing in Human-Robot Interaction. In Proceedings of the 6th International Conference on Human-Robot Interaction, Lausanne, Switzerland, 6–9 March 2011; p. 331.
- Takayama, L.; Pantofaru, C. Influences on proxemic behaviors in human-robot interaction. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2009, St. Louis, MO, USA, 10–15 October 2009.
- 42. Kendon, A. Spacing and orientation in co-present interaction. In *Development of Multimodal Interfaces: Active Listening and Synchrony;* Esposito, A., Campbell, A., Vogel, N., Hussain, C., Nijholt, A., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; pp. 1–15.
- 43. Kendon, A. Gesture: Visible Action as Utterance; Cambridge University Press: Cambridge, UK, 2005.
- 44. Yamaoka, F.; Kanda, T.; Ishiguro, H.; Hagita, N. How close? Model of proximity control for information-presenting robots. In Proceedings of the 2008 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI), Amsterdam, The Netherlands, 12–15 March 2008.
- 45. Yamaoka, F.; Kanda, T.; Ishiguro, H.; Hagita, N. Developing a model of robot behavior to identify and appropriately respond to implicit attention-shifting. In Proceedings of the 2009 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI), La Jolla, CA, USA, 11–13 March 2009; p. 133.
- 46. Yamaoka, F.; Kanda, T.; Ishiguro, H.; Hagita, N. A model of proximity control for information-presenting robots. *IEEE Trans. Robot.* **2010**, *26*, 187–195. [CrossRef]
- 47. Friedman, D.; Steed, A.; Slater, M. Spatial social behavior in second life. In *International Workshop on Intelligent Virtual Agents*; Springer: Berlin/Heidelberg, Germany, 2007; Volume 4722, pp. 252–263.
- Bailenson, J.N.; Blascovich, J.; Beall, A.C.; Loomis, J.M. Equilibrium theory revisited: Mutual gaze and personal space in virtual environments. *Presence Teleoperators Virtual Environ*. 2001, 10, 583–593. [CrossRef]
- 49. Janssen, J.H.; Bailenson, J.N.; Ijsselsteijn, W.A.; Westerink, J.H.D.M. Intimate heartbeats: Opportunities for affective communication technology. *IEEE Trans. Affect. Comput.* **2010**, *1*, 72–80. [CrossRef]
- 50. Kim, Y.; Mutlu, B. How social distance shapes human-robot interaction. *Int. J. Hum. Comput. Stud.* **2014**, 72, 783–795. [CrossRef]
- 51. Mori, M.; MacDorman, K.F.; Kageki, N. The uncanny valley. *IEEE Robot. Autom. Mag.* 2012, 19, 98–100. [CrossRef]
- 52. Bartneck, C.; Kanda, T.; Ishiguro, H.; Hagita, N. My robotic doppelgänger A critical look at the Uncanny Valley. In Proceedings of the IEEE International Workshop on Robot and Human Interactive Communication, Toyama, Japan, 27 September–2 October 2009.
- Loffler, D.; Dorrenbacher, J.; Hassenzahl, M. The uncanny valley effect in zoomorphic robots: The U-shaped relation between animal likeness and likeability. In Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction, Cambridge, UK, 23–26 March 2020; pp. 261–270.
- 54. Fong, T.; Nourbakhsh, I.; Dautenhahn, K. A survey of socially interactive robots. *Robot. Auton. Syst.* **2003**, 42, 143–166. [CrossRef]
- 55. Mehrabian, A. Nonverbal Communication; Routledge: New York, NY, USA, 1972; Volume 91.
- 56. Alibali, M.W. Gesture in Spatial Cognition: Expressing, Communicating, and Thinking About Spatial Information. *Spat. Cogn. Comput.* **2005**, *5*, 307–331. [CrossRef]
- 57. Saberi, M.; Bernardet, U.; Dipaola, S. An Architecture for Personality-based, Nonverbal Behavior in Affective Virtual Humanoid Character. *Procedia Comput. Sci.* **2014**, *41*, 204–211. [CrossRef]

- 58. Li, Z.; Jarvis, R. A multi-modal gesture recognition system in a human-robot interaction scenario. In Proceedings of the 2009 IEEE International Workshop on Robotic and Sensors Environments, Lecco, Italy, 6–7 November 2009; pp. 41–46.
- 59. Riek, L.D.; Rabinowitch, T.-C.; Bremner, P.; Pipe, A.G.; Fraser, M.; Robinson, P. Cooperative gestures: Effective signaling for humanoid robots. In Proceedings of the 2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Osaka, Japan, 2–5 March 2010; pp. 61–68.
- 60. Ge, S.S.; Wang, C.; Hang, C.C. Facial expression imitation in human robot interaction. In Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN, Munich, Germany, 1–3 August 2008.
- 61. Tinwell, A.; Grimshaw, M.; Nabi, D.A.; Williams, A. Facial expression of emotion and perception of the Uncanny Valley in virtual characters. *Comput. Human Behav.* **2011**, 27, 741–749. [CrossRef]
- 62. Kobayashi, H.; Hara, F. Facial interaction between animated 3D face robot and human beings. In Proceedings of the 1997 IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation, Orlando, FL, USA, 12–15 October 1997; pp. 3732–3737.
- 63. Breazeal, C. Toward sociable robots. Robot. Auton. Syst. 2003, 42, 167–175. [CrossRef]
- 64. Terada, K.; Takeuchi, C. Emotional Expression in Simple Line Drawings of a Robot's Face Leads to Higher Offers in the Ultimatum Game. *Front. Psychol.* **2017**, *8*. [CrossRef]
- 65. Mirnig, N.; Tan, Y.K.; Han, B.S.; Li, H.; Tscheligi, M. Screen feedback: How to overcome the expressive limitations of a social robot. In Proceedings of the 2013 IEEE RO-MAN, Gyeongju, Korea, 26–29 August 2013; pp. 348–349.
- Marti, P.; Giusti, L. A robot companion for inclusive games: A user-centred design perspective. In Proceedings of the IEEE International Conference on Robotics and Automation, Anchorage, AK, USA, 3–7 May 2010.
- 67. Gratch, J.; Okhmatovskaia, A.; Lamothe, F.; Marsella, S.; Morales, M.; Van der Werf, R.J.; Morency, L.P. Virtual rapport. In *International Workshop on Intelligent Virtual Agents*; Springer: Berlin/Heidelberg, Germany, 2006; Volume 4133, pp. 14–27.
- 68. Shapiro, A. Building a character animation system. In *Motion in Games;* Allbeck, J.M., Faloutsos, P., Eds.; Springer: Berlin/Heidelberg, Germany, 2011; Volume 7060, pp. 98–109.
- 69. Bohus, D.; Horvitz, E. Facilitating multiparty dialog with gaze, gesture, and speech. In Proceedings of the International Conference on Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction, Beijing, China, 8-12 November 2010.
- 70. Noma, T.; Zhao, L.; Badler, N.I. Design of a virtual human presenter. *IEEE Comput. Graph. Appl.* **2000**, *20*, 79–85. [CrossRef]
- Sauppé, A.; Mutlu, B. Robot deictics: How gesture and context shape referential communication. In Proceedings of the 2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Bielefeld, Germany, 3–6 March 2014; pp. 342–349.
- 72. Bremner, P.; Pipe, A.G.; Fraser, M.; Subramanian, S.; Melhuish, C. Beat gesture generation rules for human-robot interaction. In Proceedings of the IEEE International Workshop on Robot and Human Interactive Communication, Toyama, Japan, 27 September–2 October 2009.
- 73. Bremner, P.; Leonards, U. Iconic Gestures for Robot Avatars, Recognition and Integration with Speech. *Front. Psychol.* **2016**, 7. [CrossRef] [PubMed]
- Aly, A.; Tapus, A. Prosody-based adaptive metaphoric head and arm gestures synthesis in human robot interaction. In Proceedings of the 2013 16th International Conference on Advanced Robotics, ICAR 2013, Montevideo, Uruguay, 25–29 November 2013.
- 75. Hanson, D. Hanson Robotics. Available online: https://www.hansonrobotics.com/research/ (accessed on 26 October 2020).
- 76. Hanson, D. Exploring the aesthetic range for humanoid robots. In Proceedings of the ICCS/CogSci-2006 Long Symposium: Toward Social Mechanisms of Android Science, Vancouver, Canada, 26–29 July 2006; pp. 39–42.
- 77. Kätsyri, J.; de Gelder, B.; Takala, T. Virtual Faces Evoke Only a Weak Uncanny Valley Effect: An Empirical Investigation with Controlled Virtual Face Images. *Perception* **2019**, *48*, 968–991. [CrossRef] [PubMed]
- 78. Chattopadhyay, D.; MacDorman, K.F. Familiar faces rendered strange: Why inconsistent realism drives characters into the uncanny valley. *J. Vis.* **2016**, *16*. [CrossRef] [PubMed]
- 79. Argyle, M.; Cook, M.; Cramer, D. Gaze and Mutual Gaze. Br. J. Psychiatry 1994, 165, 848-850. [CrossRef]

- 80. Duncan, S.; Fiske, D.W. Face-to-Face Interaction; Routledge: Abingdon, UK, 1977.
- Admoni, H.; Scassellati, B. Social Eye Gaze in Human-Robot Interaction: A Review. J. Human Robot. Interact. 2017, 6, 25. [CrossRef]
- 82. De Hamilton, A.F.C. Gazing at me: The importance of social meaning in understanding direct-gaze cues. *Philos. Trans. R. Soc. B Biol. Sci.* **2016**, *371*. [CrossRef]
- Knapp, M.L.; Hall, J.A.; Horgan, T.G. Nonverbal Communication in Human Interaction, 8th ed.; Cengage Learning: Boston, MA, USA, 2012.
- Das, A.; Hasan, M.M. Eye gaze behavior of virtual agent in gaming environment by using artificial intelligence. In Proceedings of the 2013 International Conference on Electrical Information and Communication Technology (EICT), Khulna, Bangladesh, 13–15 February 2014; pp. 1–7.
- Muhl, C.; Nagai, Y. Does disturbance discourage people from communicating with a robot? In Proceedings of the RO-MAN 2007-The 16th IEEE International Symposium on Robot and Human Interactive Communication, Jeju, Korea, 26–29 August 2007; pp. 1137–1142.
- Mutlu, B.; Forlizzi, J.; Hodgins, J. A storytelling robot: Modeling and evaluation of human-like gaze behavior. In Proceedings of the 2006 6th IEEE-RAS International Conference on Humanoid Robots, HUMANOIDS, Genova, Italy, 4–6 December 2006.
- 87. Staudte, M.; Crocker, M. The effect of robot gaze on processing robot utterances. In Proceedings of the 31th Annual Conference of the Cognitive Science Society, Amsterdam, The Netherlands, 29 July–1 August 2009.
- Mutlu, B.; Yamaoka, F.; Kanda, T.; Ishiguro, H.; Hagita, N. Nonverbal leakage in robots: Communication of intentions through seemingly unintentional behavior. In Proceedings of the 4th ACM/IEEE International Conference on Human Robot Interaction, La Jolla, CA, USA, 9–13 March 2009; Volume 2, pp. 69–76.
- 89. Li, S.; Zhang, X. Implicit Intention Communication in Human-Robot Interaction Through Visual Behavior Studies. *IEEE Trans. Human Mach. Syst.* 2017, 47, 437–448. [CrossRef]
- 90. Sakita, K.; Ogawara, K.; Murakami, S.; Kawamura, K.; Ikeuchi, K. Flexible cooperation between human and robot by interpreting human intention from gaze information. In Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566), Sendai, Japan, 28 September–2 October 2004; Volume 1, pp. 846–851.
- Li, S.; Zhang, X. Implicit human intention inference through gaze cues for people with limited motion ability. In Proceedings of the 2014 IEEE International Conference on Mechatronics and Automation, Tianjin, China, 3–6 August 2014; pp. 257–262.
- 92. Torrey, C.; Powers, A.; Fussell, S.R.; Kiesler, S. Exploring adaptive dialogue based on a robot's awareness of human gaze and task progress. In Proceedings of the ACM/IEEE international conference on Human-robot interaction, Arlington, VA, USA, March 2007; p. 247.
- 93. Lahiri, U.; Warren, Z.; Sarkar, N. Design of a gaze-sensitive virtual social interactive system for children with autism. *IEEE Trans. Neural Syst. Rehabil. Eng.* 2011, *19*, 443–452. [CrossRef]
- 94. Bee, N.; Wagner, J.; André, E.; Vogt, T. Gaze behavior during interaction with a virtual character in interactive storytelling. In *AAMAS*, *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems*, *Toronto*, *Canada*, 10–14 May 2010; IFAAMAS: Richland, SC, USA, 2010.
- 95. Wang, N.; Gratch, J. Don't Just Stare at Me! In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Atlanta, GA, USA, 10–15 April 2010; 1241; 1249.
- 96. Kipp, M.; Gebhard, P. 'IGaze: Studying reactive gaze behavior in semi-immersive human-avatar interactions'. *IVA* **2008**, 2792, 191–199. [CrossRef]
- Shi, C.; Kanda, T.; Shimada, M.; Yamaoka, F.; Ishiguro, H.; Hagita, N. Easy development of communicative behaviors in social robots. In Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, Taipei, Taiwan, 18–22 October 2010; pp. 5302–5309.
- Crick, C.; Munz, M.; Scassellati, B. 'Robotic drumming: Synchronization in social tasks'. In Proceedings of the ROMAN—15th IEEE International Workshop on Robot and Human Interactive Communication, Hatfield, UK, 6–8 September 2006; pp. 97–102. [CrossRef]
- Lim, A.; Mizumoto, T.; Cahier, L.K.; Otsuka, T.; Takahashi, T.; Komatani, K.; Ogata, T.; Okuno, H.G. Robot musical accompaniment: Integrating audio and visual cues for real-time synchronization with a human flutist. In Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems, Taipei, Taiwan, 18–22 October 2010; pp. 1964–1969.

- Liu, C.; Rani, P.; Sarkar, N. Affective state recognition and adaptation in human-robot interaction: A design approach. In Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, China, 9–15 October 2006; pp. 3099–3106.
- Rani, P.; Sarkar, N.; Smith, C.A.; Kirby, L.D. Anxiety detecting robotic system—Towards implicit human-robot collaboration. *Robotica* 2004, 22, 85–95. [CrossRef]
- 102. Mower, E.; Feil-Seifer, D.J.; Matarić, M.J.; Narayanan, S. Investigating implicit cues for user state estimation in human-robot interaction using physiological measurements. In Proceedings of the RO-MAN 2007-The 16th IEEE International Symposium on Robot and Human Interactive Communication, Jeju, Korea, 26–29 August 2007; pp. 1125–1130.
- 103. Won, A.S.; Perone, B.; Friend, M.; Bailenson, J.N. Identifying Anxiety Through Tracked Head Movements in a Virtual Classroom. *Cyberpsychol. Behav. Soc. Netw.* **2016**, *19*, 380–387. [CrossRef] [PubMed]
- 104. Shamekhi, A.; Bickmore, T. Breathe with me: A virtual meditation coach. In *Intelligent Virtual Agents*; Springer: Cham, Switzerland, 2015; Volume 9238, pp. 279–282.
- 105. Shamekhi, A.; Bickmore, T. Breathe deep: A breath-sensitive interactive meditation coach. In Proceedings of the ACM International Conference Proceeding Series, New York, NY, USA, 21–24 May 2018; pp. 108–117.
- Park, S.; Catrambone, R. Social Facilitation Effects of Virtual Humans. *Hum. Factors J. Hum. Factors Ergon. Soc.* 2007, 49, 1054–1060. [CrossRef] [PubMed]
- 107. Stepanova, E.R.; Desnoyers-Stewart, J.; Pasquier, P.; Riecke, B.E. JeL: Breathing Together to Connect with Others and Nature. In Proceedings of the 2020 ACM Designing Interactive Systems Conference, Eindhoven, The Netherlands, 6–10 July 2020.
- 108. Gildert, N.; Millard, A.G.; Pomfret, A.; Timmis, J. The Need for Combining Implicit and Explicit Communication in Cooperative Robotic Systems. *Front. Robot. AI* **2018**, *5*, 1–6. [CrossRef]
- 109. Ley, R. The Modification of Breathing Behavior. Behav. Modif. 1999, 23, 441–479. [CrossRef] [PubMed]
- Gavish, B. Device-guided breathing in the home setting: Technology, performance and clinical outcomes. *Biol. Psychol.* 2010, *84*, 150–156. [CrossRef] [PubMed]
- 111. Li, Q.; Cao, H.; Li, Y.; Lu, Y. How do you breathe-a non-contact monitoring method using depth data. In *IEEE e-Health Networking, Applications and Services;* IEEE: Piscataway, NJ, USA, 2017; pp. 1–6.
- 112. Parati, G.; Malfatto, G.; Boarin, S.; Branzi, G.; Caldara, G.; Giglio, A.; Bilo, G.; Ongaro, G.; Alter, A.; Gavish, B.; et al. Device-Guided Paced Breathing in the Home Setting: Effects on Exercise Capacity, Pulmonary and Ventricular Function in Patients with Chronic Heart Failure: A Pilot Study. *Circ. Heart Fail.* 2008, 1, 178–183. [CrossRef]
- 113. Brown, R.P.; Gerbarg, P.L. Yoga Breathing, Meditation, and Longevity. *Ann. N. Y. Acad. Sci.* **2009**, 1172, 54–62. [CrossRef]
- 114. Khalsa, S.B.S. Treatment of Chronic Insomnia with Yoga: A Preliminary Study with Sleep? Wake Diaries. *Appl. Psychophysiol. Biofeedback* 2004, 29, 269–278. [CrossRef]
- Valmaggia, L.R.; Latif, L.; Kempton, M.J.; Rus-Calafell, M. Virtual reality in the psychological treatment for mental health problems: An systematic review of recent evidence. *Psychiatry Res.* 2016, 236, 189–195. [CrossRef]
- Faria, A.L.; Cameirão, M.S.; Couras, J.F.; Aguiar, J.R.O.; Costa, G.M.; Badia, S.B.I. Combined Cognitive-Motor Rehabilitation in Virtual Reality Improves Motor Outcomes in Chronic Stroke—A Pilot Study. *Front. Psychol.* 2018, 9, 854. [CrossRef]
- 117. Dar, S.; Lush, V.; Bernardet, U. The Virtual Human Breathing Relaxation System. In Proceedings of the 5th Experiment@International Conference (exp.at'19), Funchal, Portugal, 12–14 June 2019; pp. 276–277.

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



© 2020 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 (http://creativecommons.org/licenses/by/4.0/).