Title	Synchronized Agent Groups for Influencing Human's Movement, Synchronization, and Perception
Author(s)	Meneses, Carnero Jose Alexis
Citation	大阪大学, 2022, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/91760
rights	
Note	

Synchronized Agent Groups for Influencing Human's Movement, Synchronization, and Perception

MENESES CARNERO ALEXIS JOSE

OCTOBER 2022

Synchronized Agent Groups for Influencing Human's Movement, Synchronization, and Perception

A dissertation submitted to
THE GRADUATE SCHOOL OF ENGINEERING SCIENCE
OSAKA UNIVERSITY
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY IN ENGINEERING

BY

MENESES CARNERO ALEXIS JOSE

OCTOBER 2022

ABSTRACT

Synchronization between humans is a phenomenon that demands the fulfilment of a number of requirements, but when it occurs, it produces numerous benefits. Closeness, empathy, and coordination are some of the positive effects of synchronization. Such benefits are manifest in different environments, such as high schools, elderly care centers, entertainment parks, etc. Therefore, understanding the factors producing synchronization among humans in different environments and situations is needed that our society may increasingly experience the benefits of such effects.

Those who witness a number of synchronous movements in front of an audience are proven to be influenced by the experience and themselves caused to perform similar movements. Not only are humans able to enhance synchronization, but agents such as robots or virtual humans may also achieve this effect. Several studies have investigated the potential influence of agents on people. Android robots such as Geminoid, Erica, Da Vinci, etc. have shown effects such as body ownership or rubber hand illusion on individuals. The previously mentioned effects of androids on humans suggest that robots might be able to influence the movements of humans and their synchronization with others humans. Moreover, not only androids might be able to influence the behaviors of others. Certain small robots might be able to change the perception of robots on certain duties. As an example, CommU robots are able to increase conversation time when many CommU robots are involved in the said conversation. Due to the previously mentioned effects, this thesis explores the influence of agents with robotic and human avatars to clarify the potential influence of o agent groups on movement, synchronization, and perception.

Study One explores the influence of a projected wall group of agents on the movement and synchronization of the human hand. The influence of the projected agents on the hand movement of a human was expected to be changed by increasing the number of agents. Moreover, subjects were asked to perform a complex task to explore the synchronization between subjects and agents. The projected agents had two different avatar shapes: a robotic avatar similar to CommU, and a human avatar. In addition, the projected agents had two different speeds in their hand movement, one movement speed termed biological speed movement (Human-like speed), and the other termed linear speed movement (Robot-like speed). In other words, the experiment explored the influence of the number of projected agents, the avatar type, and speed of agents' movement on the hand's subject movement interference and the synchronization of the hand's subject. The influential factor for hand interference in the hand subject movement was found to be the number of agents and the influential factor for synchronizing the hand subject movement with the agent was the avatar type, the human one being most influential.

Study Two explores the influence of a synchronous dancing robot group on the rhythmic synchronization between two individuals and on perceived enjoyment of the interaction. In this experiment, CommU robots were controlled by two percussion instruments (electric drums) by two people facing each other. The hypothesis explored in this study was to increase the number of robots to enhance the synchronization between the two individuals. The result showed that three robots were more influential than no robots. Moreover, the enjoyment perception was enhanced more in the three-robots situation than with single robots. Similarly, the enjoyment perception was enhanced more with single robots than with zero-robots.

Study Three explores the influence of sharing the same robot body to enhance synchronization between two individuals, and enjoyment perception of the interaction. The effect of sharing two people's rhythms in robots (projecting two people's rhythms into the robots) with a single rhythm in a robot (projecting a single person's rhythm into the robots) was also evaluated. Finally, this study explores the influence of sharing musical and human beats in a robot. Synchronization and enjoyment were found to be more enhanced when humans shared the rhythm with robots than when rhythms were not shared. Moreover, the effect of sharing rhythm with multiple robots was found to be similar to that of sharing musical and human beats with a single robot.

Study Four, evaluates the influence of virtual agents on enhancing the emotional perception of joy and emotional conveyance of joy of a person on video-conferencing systems. In order to determine the influence on emotion perception and conveyance, the use of one virtual agent, two synchronous virtual agents, and no virtual agents was proposed. One virtual agent was found to enhance the emotional perception of a human more than no virtual agent. Further, two virtual agents were found to enhance the emotional perception and emotional conveyance of a human more than a single virtual agent.

This thesis establishes the fact that the influence of synchronized agent groups increases when the number of agents in the group increases. Synchronized agent groups were proven to influence the hand movement of people, the synchronization of rhythmic movement, and the perception of the joy of people.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Synchronization

Synchronization is the process of adjusting the rhythms of oscillating objects via their weak interactions [1]. The synchronization process can appear in objects or in animals. An example of this non-human synchronization is illustrated on the figure 1.1. These type of phenomenon could be refered as a frequency locking.

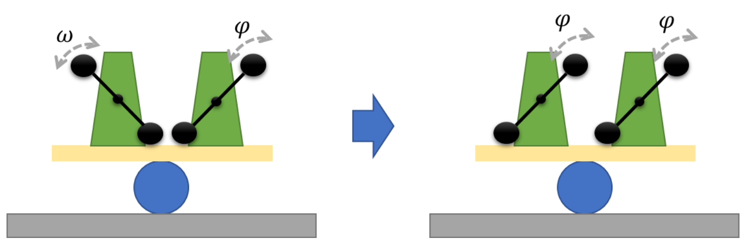

Figure 1.1: Two metronomes can start with two different phases ω and ϕ and synchronize their phases after some time, ending with the same ω phase.

There is a difference between the synchronization appearing in objects than in humans. Humans unconsciously synchronize their body behavior by sharing rhythms [2] and coordinating their actions [3]. This type of synchronization usually appears as entrainment and could be described as Rhythm sharing. An illustrative example of this synchronization is shown on the figure 1.2. Rhythm-sharing is found in various mechanical activities, like rocking in chairs [4], walking [5], balancing a pendulum [6], and paced leg-stretching [7]. Synchronization improves social relations between humans. Mothers develop a stronger feeling of closeness to babies who imitate them [8]. Synchronization between people generates positive impressions about each other and their fostered cooperation [9], and enhances empathy [10]. In human–robot interaction, synchronization provides a more pleasant and a smoother experience with the robot [11]. As demonstrated in the above studies, inducing synchronization with humans contributes to improving social relations.

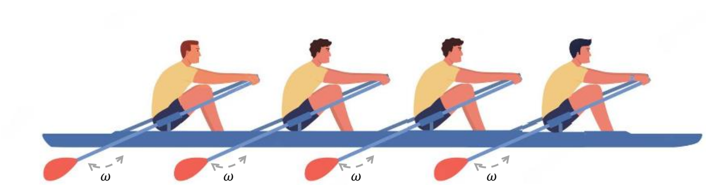

Figure 1.2: Synchronization in humans by entrainment. Rowers can make other rowers be synchronized (having same ω) by entrainment.

Humans have created and experienced music for more than 36,000 years [12], and they tend to synchronize their behaviors with music [13]. Moreover, music evolved as a way of synchronizing individuals and groups while promoting social bonds and group cohesion [14, 15, 16]. It has been reported that people in a group tend to synchronize their clapping [17]. Additionally, when two or more people tap their fingers, they tend to synchronize [2]. Furthermore, children synchronize their body movements when they walk with music playing in the background [18]. In the same way, children synchronize their movements to a robot when it makes synchronous movement to music [19] and children synchronize their movements more when a robot dances to the music [20]. In addition, when a robot's dance movement synchronizes with the music, it enhances the perception of lifelikeness, dance quality, and entertainment level of the robot [21]. At the same time, people who see a robot synchronize with a song perceived the song to be more likeable [22]. Furthermore, robots are able to synchronize with a person while the robot and the person are playing drums [23], synchronize with a person more naturally while playing drums using turn taking methods [24], or synchronize with a group of people while they are playing drums [25]. People need a synchronization process [26], and music offers a cost-effective medium to influence human rhythm and synchronization.

Humans can synchronize with robots in non-musical environments as well. For example, humans can synchronize their eyeblink timing to androids when they talk with them [27]. Similarly, humans synchronize the eye blink timing with the timing of the robot's movements and follow them when there is inconsistency with the humans' movements [28, 29]. Moreover, robots can synchronize their movements with a group of humans to improve the social relations with the other group of humans. For example, robots can synchronize their drumming time to the human's drumming timing in order to help the human to improve their drumming performance [30]. Moreover, synchronizing the gesture expressions of a robot with a human's nodding behavior can enhance the communication between the human and the robot [31]. Therefore, the robots' synchronization with humans might be useful for more social tasks.

With the "rubber hand experiment", humans confuse an artificial rubber hand with their real hand during multimodal synchronization between the visual stimulation of the rubber hand and the tactile stimulation of their hand [32, 33]. A similar illusion is induced when a robot hand provides a subject the experience of observing synchronized movement of the robot hand with subject's own movement [34] or when the robot uses subject's brain signal to perform body movement [35]. Other studies show that humans unconsciously produce muscle activity corresponding to the movement produced by the confused body part, depending on the degree of confusion [36]. Therefore, the body movements of robots are more influential to a subject's rhythm and social relationship with others when the robot's body parts follow the subject's rhythmic movement. These findings suggest that robots that move in-sync with a subject are more influential in terms of the natural synchronization of the subject than out-of-sync robots. In this paper, it was explored the effect of robot's movement being not only in-sync with the rhythm of another person in front of the subject but also with the subject.

1.2 Number of agents

Many people are known to influence individual behavior so that people perform tasks they would not normally do [37]. Moreover, people take riskier decisions when they belong to a group than when they are alone [38]. Political parties influence members' beliefs and persuade them to accept party policies [39]. Similarly, a group of humans can influence a human outside of the group to choose an obviously incorrect answer for an easy question, by intentionally showing that all members of the group chose the same incorrect answer for the question; this is an example of a phenomenon known as social pressure [40]. Moreover, it is known that groups of people influence members of their groups in their behavior and perception of others [41]. In an experimental survival task, if an autonomous robot is included in the group, it can influence how the humans solve tasks [42]. Additionally, a human in control of a robot can be influenced by other robots to behave more ethically towards the robot. [43]. Regarding interactions with robots, humans follow a conversation more easily when there are many robots instead of one [44, 45]. These findings suggest that humans are more influenced by groups than by a single agent, even if part of the group consists of robots.

Groups of agents are also influential in terms of group perceptions regarding decisions such as punishing others even when they don't want to do it [46]. A group of robots can make a social pressure to a human and lead him/her to choose an incorrect answer to an easy question [47]. Similarly, compared to a single robot, a group of robots can enhance the robustness of a conversation, which causes to prolong the time of the conversation [48, 49, 45]. On the other hand, some research use robots as a member of a human group in order to influence the behavior of humans outside of the group. For example, it was shown that patients feel more secure when an android robot helps a medical doctor to diagnose [50]. Moreover, the impression of an interviewee was evaluated as more possitive by an interviewer when an android has eye contact with the interviewer for giving possitive feedbakcs about the interviewee [51]. Further, there is evidence that groups of robots can synchronize with humans in musical environments [52]. Synchronization is well-known to enhance empathy and feelings of closeness to others [53, 54]. Thus, the influence of a group of synchronized agents can be greater than the influence of a single agent.

1.3 Interference

The development of novel communication technologies, e.g., humanoid robots [55, 56, 57] and virtual reality avatars [58, 59, 60], influence social behavior in society [61]. A type of influential social behavior was shown in human kinematic performance which can be influenced by observing an incongruent movement of a different human [28], naming this type of influence "interference". Analyzing the potential factors that produce this interference is fundamental to understanding the effect of emerging technologies on humans. The interference effect has been studied under different conditions, including observing humans on screens, robots, and in virtual reality [62, 63]. The present chapter explores the influence of a group of agents, as the type of movement of the agent and the virtual agents' avatars.

Virtual avatars have also been shown to influence humans shopping decisions and behaviors in the virtual world depending on the avatar shape [64, 64]. Moreover, previous studies reported that the shape of a virtual avatar changes the perception of someone's weight [65] and influences the Kinematics movement interference [66]. Not only the shape, but the size of the avatar may influence the perception of the avatar [67]. Also, and importantly, the movement of the avatar may also contribute to a perception of the avatar as being more realistic [68]. Thus, finding appropriate embodiment avatars plays a crucial role in increasing motion interference.

Timing action synchronization is a feature that humans can naturally develop, e.g. humans naturally synchronize their claps in environments where there is a group of humans clapping [17] Even though humans can naturally develop synchronization, not always as easy to synchronize with others, e.g. humans that have problems of concentrating or can not follow the beat in music [69, 70]. Synchronization in different environments depends on the task the user is performing e.g. the muscles of humans synchronize their performance depending [71]. Some actions, especially in a game, request coordination that doesn't require having the same starting timing movement [72]. This complementary synchronization depends on the phase of the agents' time, meaning the more the phase are kept constant but not the same, the better performance they have [73]. To explore the influence of the virtual agents on different types of synchronization, it was proposed to use antiphase synchronization movements, defining it as the ninety-degree phase deviation.

1.4 Emotional modalities

There are not enough good communication modalities for helping the emotional perception of humans in current technologies e.g., Video conference systems or virtual seminars. There are many benefits on finding good communication modalities in emotional perception, e.g., it has been reported that correct emotion perception improves the communication skills of individuals [74], enhances leadership [75], and helps create harmonious social interactions [76, 77]. Humans can naturally recognize emotions and it is known that humans can perceive basic universal emotions across different cultures [78]. Due to the evidence mentioned, finding good modalities for an appropriate emotional perception of the interlocutor in current technologies might help to have accurate communication. Emotional conveyance is as important as emotion perception and can be defined as the ability to accurately express emotions. Emotion conveyance improves and maintains social relations by communicating ideas precisely [79, 80, 81, 82]. Emotion conveyance also helps to have a good degree of empathy between peers [83]. In a previous work, people who used a gestural modality i.e., smile, conveys positive feelings such as trust for cooperation [84]. This gestural modality is not enough for conveying an emotion in videoconferencing due to video problems or bad camera angles. Therefore, finding good modalities for conveying emotions (i.e., smiling) is important in current technologies for enhancing communication. In particular, humans use different types of modalities for conveying and perceive emotions such as voice pitch [85], facial expressions [86], body postures [87], context information about the environment a person is experiencing [88], intergroup information such as race [89], as well as the combination of all of the above [90]. In current technologies, the most common modalities used are voice pitch, environment information, and facial expressions [91]; however, as stated before, these modalities might need extra modalities for having more accurate emotional conveyance and perception. Therefore, it might be beneficial to add virtual agents as a modality which contains body postures as a method for improving the emotional perception and conveyance. The use of virtual agents as a modality for communication is on the rise. For example, in a virtual reality environment, people can recognize emotions in virtual character agents, similar to human faces [92]. Another study combined human facial expressions and body postures on static images, showing that when face expression and body posture have a congruency with the emotion, the emotion is perceived more precisely [93]. However, these studies lack an evaluation of the perception of emotions over videos for using these agents for helping humans. In this thesis, it was proposed to use face and body as they are the easier modalities to recognize in a virtual environment which has other modalities at the same time. As the space for the virtual agent is reduced, it was proposed the use a virtual agent that has a small body and a larger human face as an extra type of modality. Selecting the larger face is optimal as the face contains more information about the emotions. It was expected that these virtual agents will help humans to perceive and convey their emotions easier. As happy emotions are easy to understand for humans, it was proposed to use happy expression on the agents for perceiving and conveying a not obviously happy person. Moreover, it might be possible that a single virtual agent might not be able to enhance the perception and conveyance of emotions for having not enough strong modalities. Due to this, it was proposed to use two synchronous virtual agents. It was proposed to use synchronization because synchronous groups lead to better coordination [94], understanding [95], cooperation [9], and rapport perception [96]. Moreover, previous studies have shown that the more synchronized agents there are, the more influence they have on participants that are not part of the group [52]. In addition, it is known that group of happy faces under laughter enhances the emotional perception [97]. Therefore, it is worthy to increase the number of a single agent to two synchronous virtual agents for having a group of agents that might be more influential. The virtual agents' similarity can enhance the degree of perception and conveyance of the emotions that one person is showing due to their avatar characteristics in a group.

Therefore, this research focused on exploring the influence of synchronized agents on the synchronization of humans in terms of body movements and facial expressions, by building systems controlling the number, movement, and synchronization of the agents. The research conducted for this aim were shown in fig 1.3. In order to explore the body movement characteristics that influence the synchronization with a single person, the effects of the number of agents, the type of agent, and the speed of the movement were investigated to influence the synchronization of the movement of the hand of a human with the agents. In other words, the influence on the synchronization of a single human moving his/her arm with a group of projected agents (first study) was studied. Similarly, in order to explore the body movement characteristics that influence the synchronization between two people, the effect of the number of robots was investigated to influence the synchronization of rhythmic movements between two humans. In other words, the influence on the synchronization between two humans' rhythmic movements (second study) was studied. Also, in order to explore the body movement characteristics that enhance the influence on the synchronization between two people, the effect of sharing or not sharing the control of the body of a group of robots was investigated for enhancing the synchronization of rhythmic movements between two humans. In other words, the representation of rhythms by sharing the body of a group of robots (third study) was studied. Additionally, in order to explore the facial expressions characteristics that influence emotion perception, emotion conveyance, and emotion visual perception of a person, the effect of the number of agents was investigated to enhance the facial expression of a person. In other words, the influence of the number of agents for enhancing facial expression (fourth study) was studied.

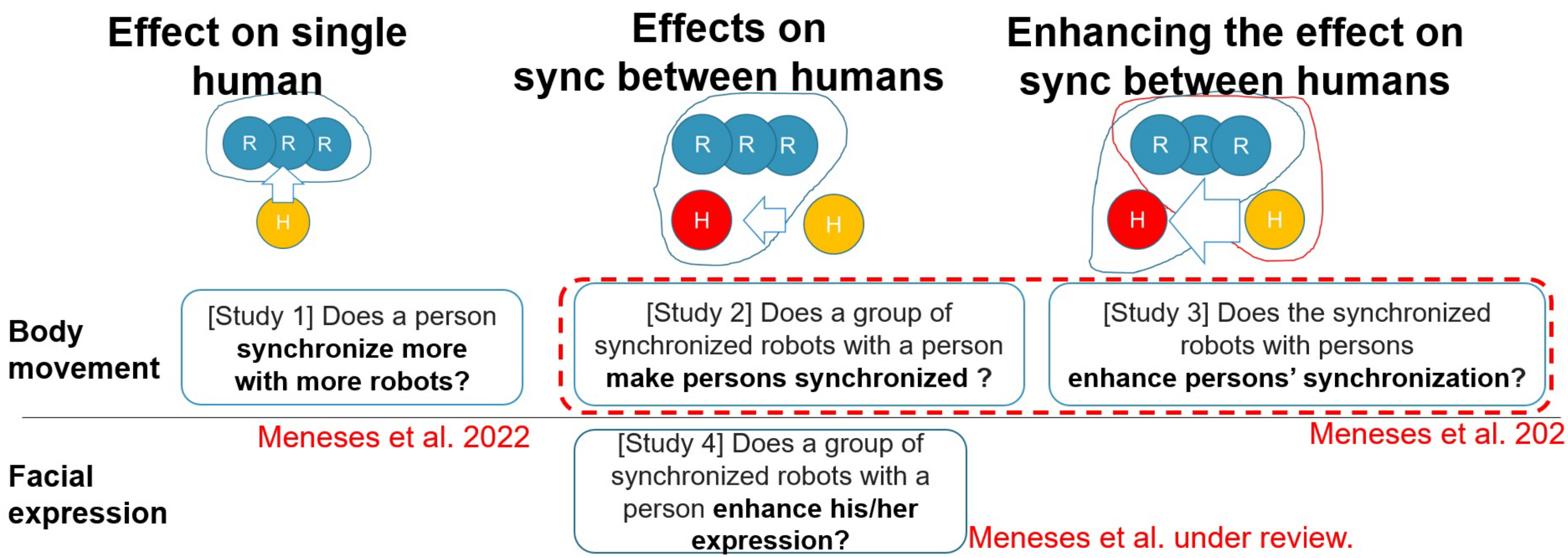

Figure 1.3: Research map of this study.

CHAPTER 2

RELATED WORKS

2.1 Interference

Movement interference is a phenomenon that has previously been studied using humans or robots. However, these studies did not investigate the effects of the interference of groups (number of agents in groups) and the type of agents (human or robot). The movement interference appears in many degrees, as William James described that imaging on performing a movement awakens the possible movement to some degrees [98] In other studies, it was verified that imaging an action, observing someone's action, or trying to represent the action on person self mind excites the muscles to be used to execute the action imagined, observed or tried to be represented [99, 100]. This phenomenon leads to the assumption that the action performance of a person may be able to be influenced by others humans' performance. For example, a person's hand's muscle potential is linked to the observation of an active movement while being the observer of a different person remained linked temporally to the observed movement [101]. In addition, the timing of the muscle initiation of a person becomes slower after watching another person moving their finger in similar patterns, or even when the other person was grasping an object in a similar way to the person [102, 103]. Moreover, the actions of a person are another characteristic despite the time. Wilner found that people tend to imitate models' actions when they perform a synchronous action [104]. In this study, he asked the subjects to move their pointer finger on a table with points at a certain speed while watching a video of a human hand moving. He found that the imitation accuracy on time increased when the cognitive demand of the action was less intensive. In his study, he suggested that non-goal actions are moreover to be influential on visuomotor mapping (interference movement) in contrast to goal action which is moreover influential on accuracy. Hayes extended Wilner's work by creating a different task with mouse pointers [105] Hayes drow two red points on a monitor screen and show videos before the experiment starts where the mouse cursor moves from one point to the other one. He manipulates the videos for showing atypical movement speeds on the cursors, assuming that when the cursor moves more human-like, the cursors will be able to move similarly to the human-like movements. He found that in cursor movements, the people were more willing to learn the imitation the movement when the movement was more human-like. None of these studies explore whether a group of agents may be able to influence the action of the agent. Moreover, the evaluation of a high cognitive demand task was not enough studied. Studying a task such as trying the people to keep asynchronous movement may clarify the human movement interference. The movement interference has been studied on avatars as well. Kupferberg found that the movement interference depends on the similarity of the movement that the agents have [106]. In this study, she prepared a robotic arm that moves in front of a human with orthogonal and parallel movements. They used 4 stimuli: a human, a humanoid torso, a robotic hand with a table base, and a robotic hand with the base attached to the wall. They suggest that the configuration of the speed is more important than the presence of human-like features, like body shape. In this study, they also suggest that the biological movement of the hand as a human may not influence the interference. In opposition, Chamide [29] suggested that humanoid robots (without faces or any social stimuli) may be more influential for having anthropomorphic body shapes. He found that human-like movement speed was significantly more influential than robot-like movement speed. A similar conclusion was proposed by Gandolfo [107] who compared a virtual humanoid agent with a human-like body shape, a real human, and a non-humanoid agent. The task was to grasp an object instead of moving a hand. They compared the position variance of the hand at grasping finding no significant difference in the results. None of these studies, try to compare the social aspects of the avatars, they focus on the movement and they did not consider whether the conception of the avatar would influence the movement of the subject to be evaluated.

2.2 Emotional modalities

In this study, it was discussed that visual clues in communication might alter a person's perception of the message. Park[108] examined the effect of synchronous text on conveying emotions. The researchers compared a text synchronized with the speed of the writer and a text presented at once by recording the keys pressed to create the synchrony condition. The study found that using synchronous text to display the communication made the communication feel more emotional, which led the subjects to communicate more effectively. In other words, there was better communication when subjects correctly recognize others' emotions. Moreover, Ziembowicz [109] proposed that having a chat displayed in rhythm and tempo of the produced text help to identify the emotional states of users. They found that people are better at recognizing the emotions in the text the more they used this synchronous chat. But text modality was not the only visual clue able to enhance the emotion conveyance. Whitaker and Connail[110] found that people feel more comfortable using video conference system than using only audio conference systems, suggesting that using video enhance the emotion recognition of others. Moreover, they found that there are two main classes of visible information named them visual behaviors of the participants and the visible environment. Previous studies have focused their effort on demonstrating the capabilities of video communication. Bailenson[111] found that the mutual gaze in virtual avatars enhances social inhibition, inducing the subjects to perform worse than when they were not being watched by the virtual avatars. Bickmore[112] found that the immediacy empathy frame can be enhanced by using zooming on the face of an avatar, focusing the avatar only on the face. Kang [113] found that avatar videos helped the social richness of the interactions. Avatars have been studied to enhance communication in different aspects. Nass and Moon [114] found that creating an effective avatar that affects humans' emotion perception factors such as stereotypes of avatars or classifying the computer by ethnically identifying with computer agents influences the emotion perception. Moreover, humans can express their politeness and reciprocity towards agents. Finally, they showed that providing a specialist set of avatars on the television set in contrast with a general set of avatars enhances the perception of the content due to the emotional state of the audience. Sproull [115] investigated the responses of people to a face that produces voice by synthesizing a talking. In their study, they proposed that adding natural texture to the human avatar would be able to interact with people. Even when the voice was not natural, the face itself was able to enhance feelings like happiness, likability, trustworthiness, and others. This study shows the importance of having an expressive face for enhancing the emotions recognized by the people. Kopp [116] also studied the effect of a virtual avatar named Max on a museum to express more emotions and enhance the emotion recognition of the agent.

In this study, they focus on expressing the greetings and the words correctly to the meaning. The main objective was to express emotional information through the agent. They were able to imitate the face-to-face conversation with the max agent. None of these studies focus on checking whether an avatar of a group of avatars may be able to convey emotions and help others to identify the emotion of a person speaking with additional avatars attached to them.

CHAPTER 3

STUDY I: INFLUENCE OF AGENTS IN MOVEMENT INTERFERENCE AND SYNCHRONIZATION

The interference effect has not previously been studied using a group of synchronized agents or verityting the type of avatar influence. Adding the group of agents in virtual reality environments may influence the result of the movement interference in groups. Moreover, these agents might be able to enhance the synchronization timing of the movement of the hand. Exploring the influence of an avatar's characteristics on a group of synchronized agents will help to design more influential systems of virtual agents in the future. In this chapter, two different projection groups of avatars were used to increase motion interference and synchronization.

3.1 Experiments

3.1.1 Methods

To investigate the effect of multiple agents on movement interference, it was conducted an experiment in which it was investigated combinations of the following variables: Number of agents (one, two, and three), type of avatar (human and robot), and agent behavior (human-like agent moving with biological movement and robot-like agent moving with linear movement). In the study, the agents moved their hands orthogonally to the movement of the subjects to influence the subjects' movements. Each group of variables was separated depending on the subject movement i.e., vertical hand movement or horizontal hand movement. In this experiment, it was hypothesize that (H1) the higher number of agents, the higher movement interference they influenced, (H2) the human agent interfered more than the robot agent the movement, and (H3) the biological moving agent interfered more than the linear moving agent. It was additionally hypothesize that (H4) the higher the number of agents the higher influence on the synchronization between the virtual agents and a person, (H5) the human agent influence the synchronization between the projected virtual agents and a person more than the robot agent, and (H6) the biological movement agent influence the synchronization between the projected virtual agents and a person more than the linear moving agent. The subjects provided written consent, and the experiments were approved by the ethical committee.

3.1.2 Subjects

Twenty-four subjects participated in all the conditions (13 females and 11 males, mean age=22.17, standard deviation=3.69). The order of conditions was counterbalanced for the number of agents, type of avatar, and agent behavior. After the orthogonal movement, was prepare the synchronization movement conditions.

3.1.3 System

For this study, it was developed a system for tracking hand movement. The system used two Optitrack V120 trio desktop trackers for recording the position of passive infrared markers with a frequency of 120Hz. The Optitrack cameras were placed side by side with the distance of 3 meters from each other and put on a cylinder bar with the height of 1.8m. The trio camera had 3 camera sensors, with a shutter speed of 1ms, a focal lens of 3.5mm, and a low pass filter of 800nm IR. The measurement error of the hand position of the participants was around 0.8% based on the frequency of the cameras. Two client programs were prepared to record every point after synchronizing their timing by using network time synchronization. As the sensors did not have an NTP server, it was needed to estimate the package delay after some seconds of tracking the points to be able to recover the possible missing positions due to the loss of focus of the markers in front of the cameras. It was tracked the first 10 packages of each Optictrack, calculate the time difference between them and calculate the time average for having a time delay between the cameras. This could be explained with the following formula:

$$\gamma = \frac{\sum_{i=1}^{N}(\mu_i - \zeta_i)}{N} \tag{3.1}$$

where γ is the average delay time, μ_i is the delay time between tracked point on the Optitrack1 and ζ_j is the delay time between tracked point on the Optitrack2 and N is the number of packages of each Optitrack.

The average delay time was used for calculating the difference between the receiving time and the estimated real-time measurement. This estimated real-time measurement was stored on a variable in the program before the experiment started. The starting recording of the tracking point was received by an additional control socket which was developed by using socket.io on a Nodejs server (Figure 3.1).

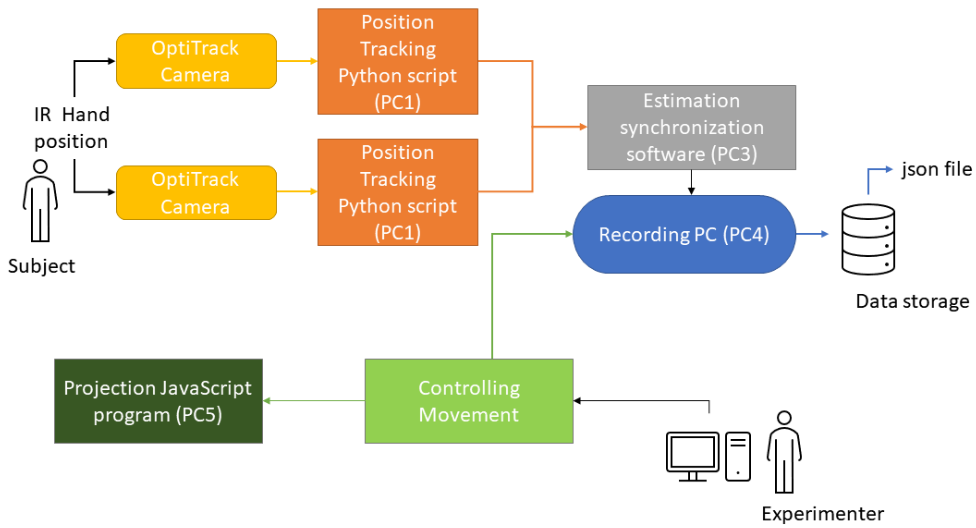

Figure 3.1: System block diagram implemented for recording hand position movement.

The robot model was displayed by using the three.js library. The library allows moving the robot remotely by using WebSockets. An additional signaling server was created to receive the command from the experimenter to start the condition required for the experiment. After the experiment finished, there was a function created to send a message to the recording program to store the data on a JSON file by using Nodejs information. The algorithm of the system is shown on Figure 3.2.

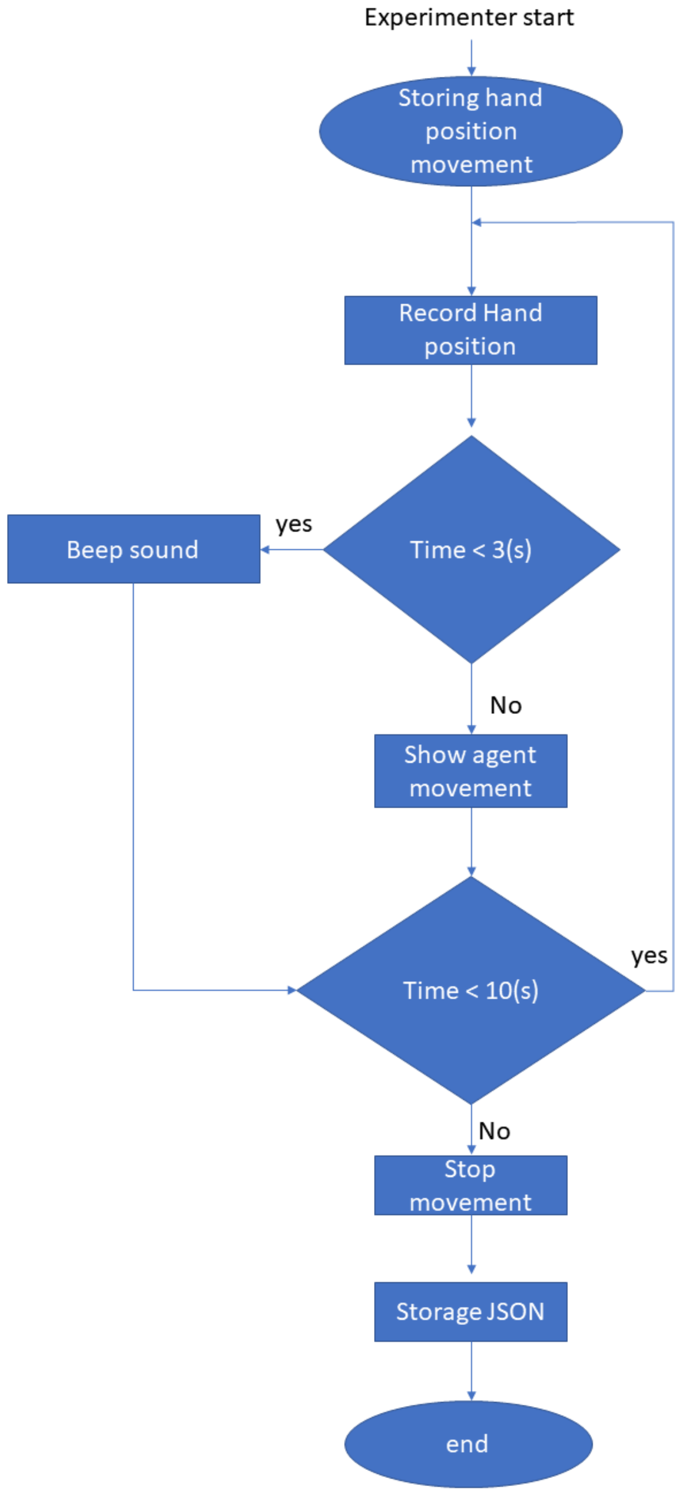

Figure 3.2: Algorithm system flow of the system for showing agent movement.

As the robot needed to have biological movement, it was recorded a human model movement to simulate the biological movement in the robot. The robots were able to move their hands in synchronization due to the time attribute javascript scope shared between the robot model displayed on the screen system.

3.1.4 Experimental setup

Two Optitrack cameras were placed one in front of the other on cylinder bars facing the passive retroreflective marker. A chair was placed 1.2m away from the wall. Before the experiment, a background image was projected on the wall. The projection image had a resolution of 1280x720 pixels and a size of 2m in width and 1.125m in height. The environment set-up is shown in Figure 3.3 . The position of the marker was recorded using a Python script that connected through TCP protocol to the Optitrack software, and stored the movement of the tracker on two separate files.

Figure 3.3: A person standing in front of the projection.

3.1.5 Avatar of the agents

One colleague was recorded moving his hand for projecting a human avatar on the wall. The size of the human avatar was 80% of the height of the projection window on the wall (0.9m) and 0.3m in width. Figure 3.4, shows an example of the projection of the human avatar. The human avatar was made using a video of a human and edited frame by frame to create a constant speed naming it linear movement.

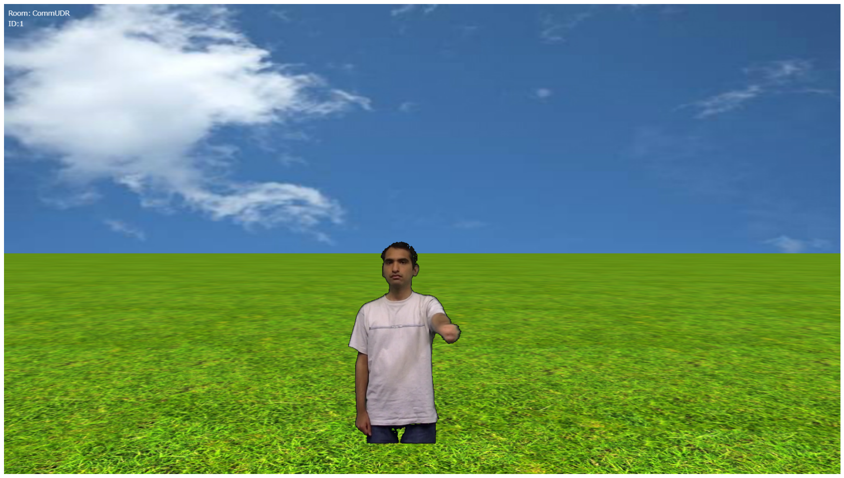

Figure 3.4: Human avatar projected in the experiment.

A robotic model was used to project the robot avatar on the wall. The size of the robot was similar to the human avatar. The robot was a humanoid robot based on a robot named “CommU”, developed by Vstonre Co., Ltd., in collaboration with Osaka University. The projected robot head employed 3 degrees of freedom (DoF), two eyes with 3 DoF, upper eyelids with 1DoF, a mouth with 1DoF, two arms, each with 2 DoF, and a waist with 2DoF. This model was manipulated using Three.js and WebSockets. Figure 3.5, shows an example of the projection of the robot avatar. The movement of the robot was tracked to a human avatar for creating the agent biological movements on the robot.

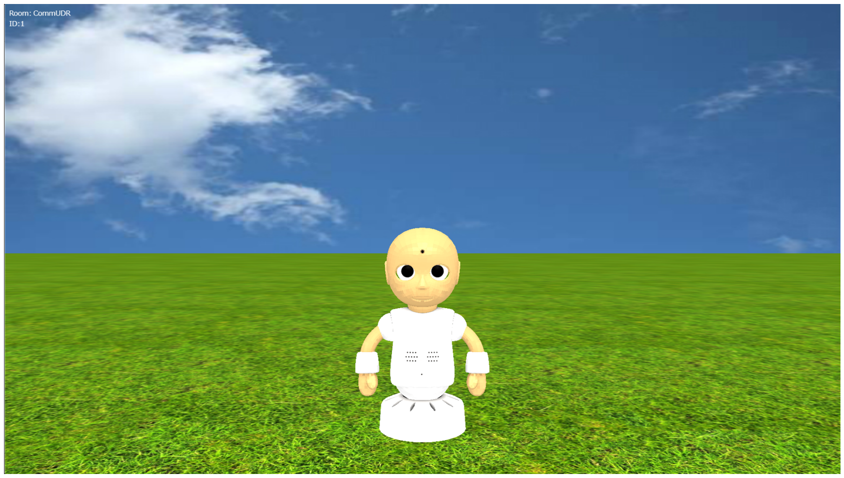

Figure 3.5: Robot avatar projected in the experiment.

The controlling of the avatar movements was adjusted to start at same time and end at same time. The human avatar was able to move with the same speed of the robotic avatar for biological and linear movement.

3.1.6 Procedure

All instructions, including informed consent, were provided via written documentation. The subjects' tasks were explained before the experiment. The subjects were instructed to move their hands vertically and horizontally on hearing a metronome at 2Hz (120bpm) before the experiment started. The trained session consisted of 20 sessions (5 seconds for each session) with a total training time of 5 minutes. Before each set-up in the experiment, the subject was instructed to move their hands in vertical or horizontal directions. An intermittent beep sound was played at 2Hz for 3 seconds at the beginning of each session, muting it after the 3 seconds. The subjects were instructed to start their hand movements after hearing two beeps, asking them to synchronize their movement with the third beep. The projection images were displayed after the sixth beep. At the end of the experiment, a beep sound was played to make the subjects stop. Subjects moved their hands for 15 seconds in 24 conditions: 2 type of avatar (robot/human), 2 types of movement (Biological movement, linear movement), number of agents (one/two/three) and 2 movement direction (Vertical/Horizontal). Additionally, the subjects attend to other 24 conditions on movements corresponding to the parallel movement of the hand. The subjects started on opposite movement times by adding an extra beep to the 3-second initial condition beeps and having the same independent variables of the kinematics movements. The movement of the hand started on a 90-degree difference in phase, due to this, the movement was named antiphase. Following each set-up, subjects were allowed to rest for 15 seconds. If the subject felt fatigued, the experimenter stopped the experiment for 5 minutes to allow the subject to recover from fatigue. It was asked one random question regarding physical feeling in the subjects' hands at the end of every set-up. Figure 3.6, shows an example of the movement recorded by the subject.

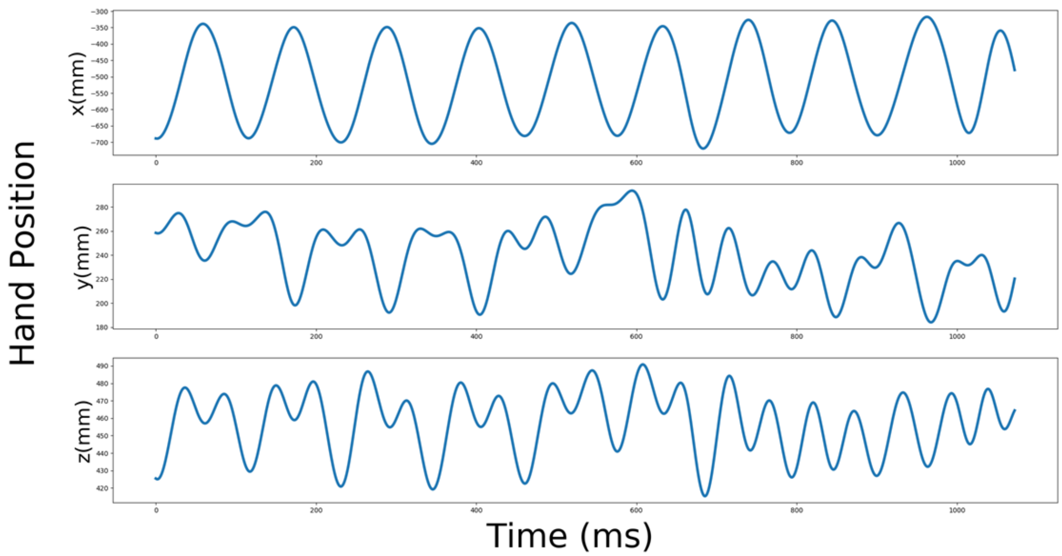

Figure 3.6: Recorded movement of the subject hand position. The top graph shows the position of the hand on the x-axis in time(ms), the middle graph shows the position of the hand on the y-axis in time(ms) and the bottom graph shows the position of the hand in the y axis in time(ms).

As the figure 3.7 shows, the signals obtained from the cameras were combined for compensating possible oclussion. In the combined signal, it was calculated the standard deviation of the orthogonal axis.

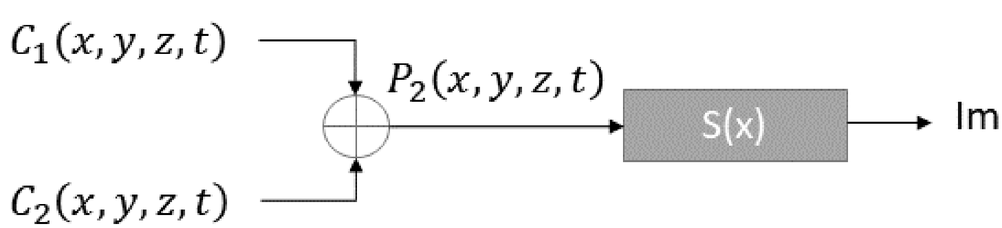

Figure 3.7: Process of combining data, calculating the standard deviation S(x) as formula 3.2 shows, where x is the orthogonal axis and Im is the interference movement index.

The data obtained of the position of the hand was processed after the experiment finished. The orthogonal plane of the position of the agent hand was evaluated for the interference movement. In the case of the antiphase synchronization, the parallel plane of the position of the agent hand was evaluated.

3.1.7 Evaluation on interference movement

It was evaluated and compared the kinematics interference movement of the subject hand in each condition. As it was utilized two cameras to avoid loss, it was combined the signals and processed them to have the first 14 movements of the hand as it was assumed that fatigue developed after 7 seconds of moving the hand due to the occlusion of the markers. It was processed the incongruent axis as the formula 3.2 shown.

$$\sigma_T = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N}} \tag{3.2}$$

where σ_T is the standard deviation of the position of the incongruent axis, x_i is the position of the incongruent axis on time i, μ is the mean of the position of the incongruent axis and N is the number of sample analyzed.

In other words, the interference movement was the standard deviation of the incongruent axis of the position of the marker of the hand.

3.1.8 Evaluation on Synchronization

Evaluating the antiphase synchronization in time requires a measurement based on the phase changes. As the interested was focused in understanding the changes in the synchronization, project the cycle starting time and calculate the cycle phase value of an expected agent's trajectory. By calculating the cycle phase changes, it was evaluated the variation of the synchronization between agent and subject. The changes in the synchronization based on the peaks of the subject trajectory hand position was evaluated. As it was utilized two cameras to avoid loss, it was necessary to combine the signals of both cameras. it was processed to have the first 14 movements of the hand as it was assumed fatigue was developed after 7 seconds of moving the hand, due to the occlusion of the passive marker placed on the subject.

A sinusoidal wave that started on the antiphase movement of the subject's first movement was made (ninety degrees phase) and it was named as the expected trajectory of the agent. To calculate the synchronization between the subject and the agent, the peak values of the trajectory subject's hand position into the expected trajectory agent's hand position were projected so that the first peak projection was in the minus ninety-degree phase of the trajectory agent's hand position. In order to obtain the phase value for each correspondent agent peak on the periodic movement of the subject's hand, the inverse sinusoidal function of the subject points that time correspond the agent peak was calculated as the formula 3.3 shows. Figure 3.8 shows an example of standardized subject's hand trajectory and the projection of the peaks on the y plane or x plane for the vertical or horizontal movement respectively.

$$\theta_s = \sin^{-1}(P(t)) \tag{3.3}$$

where θ_s is the phase of Subject in agent movement peaks, P(t) is the position of the hand subject at time t and t is the time of the agent phase position at ninety degrees.

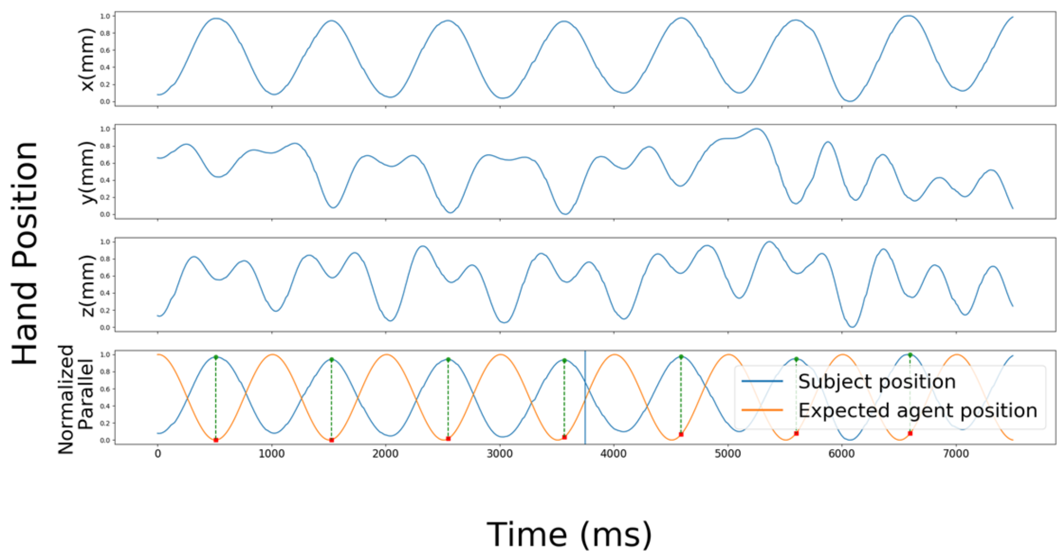

Figure 3.8: Recorded movement of a subject hand position and expected movement of the agent. The top three graphs show the standardized hand positions on the x, y, and z-axis, respectively. In the fourth graph, the standardized hand position on the x-axis (the same as the one in the first graph) is shown as a blue trajectory while the expected agent hand position is shown as an orange one. Circles indicate the peaks of the trajectory of the standardized hand position while the squares indicate their projected points on the trajectory of the expected agent position. The fourth graph shows the subject's hand trajectory in which the first pick projection phase value is minus ninety degrees difference.

The signals was divided into two sections by separating the phases projection of the first half time and the latter half time as the formula 3.4 shows.

As the figure 3.9 shows, the signals obtained from the cameras were combined and the synchronization index of the parallel axis was calculated.

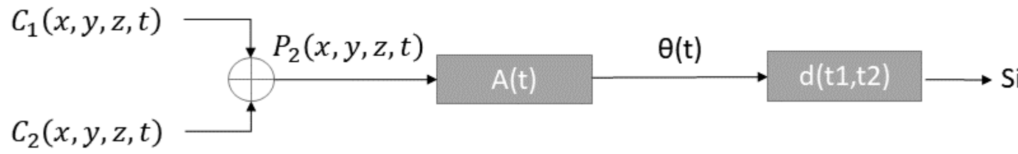

Figure 3.9: Process of the combining data, calculation the agent peaks hand positions time into the subject agent phase (Formula 3.3) and calculating the difference between the average of phases (formula 3.4).

$$Si = \frac{\sum_{i=1}^{m} \gamma_i}{m} - \frac{\sum_{i=1}^{n} \theta_i}{n} \qquad (3.4)$$

where θ_i is the phase on the prior half time i, n is the number of phases corresponding to peaks in the prior half, γ_j is the phase on the latter half time, m is the number of phases corresponding to peaks in the latter half and Si is the Synchronization index.

The difference between the average of the phases in the first half and the latter half was the synchronization index.

3.2 Results

In the following result report, M means mean, SD means standard deviation and p is the calculated probability of the alpha of the utilized t-test.

3.2.1 Results on movements interference

A three-way ANOVA for the horizontal movement interference was conducted, the results of which are shown in Table 3.1. There was not found any significant interaction effect on the variables. On the other hand, a main effect ($F(2, 276) = 3.72$, $p<.05$) in the number of robots was found. Then, post-hoc analysis utilizing Bonferroni correction (here, adjusted alpha levels of 0.017) was conducted (Figure 3.10). The results show that three agent set-up has greater interference in the movement (M=33.38 mm, SD=19.15 mm) than the one agent set-up (M=27.3 mm, SD=14.86 mm), where $t(190) = 2.46$, $p=.015$, Cohen's d=0.35. In contrast, the two agent set-up (M=28.46 mm, SD=13.98 mm) was not significantly different from the one agent set-up (M=27.3 mm, SD=14.86 mm), where $t(190) = 0.56$, $p=.578$, Cohen's d=0.08; similarly, the three agent set-up (M=33.38 mm, SD=19.15 mm) was not significantly different from the two agent set-up (M=28.46 mm, SD=13.98 mm), where $t(190) = 2.03$, $p=.043$, Cohen's d=0.29.

Table 3.1: Three-way ANOVA for horizontal interference movement.

	df	sum_sq	mean_sq	F	PR(>F)
C(Number of Agents)	2	2003.79	1001.89	3.72	0.03
C(Type of avatar)	1	12.56	12.56	0.05	0.83
C(Type of movement)	1	33.69	33.69	0.13	0.72
C(Number of Agents):C(Type of avatar)	2	15.27	7.64	0.03	0.97
C(Number of Agents):C(Type of movement)	2	7.14	3.57	0.01	0.99
C(Type of avatar):C(Type of movement)	1	3.56	3.56	0.01	0.91
C(Number of Agents):C(Type of avatar):C(Type of movement)	2	61.87	30.93	0.11	0.89
Residual	276	74246.1	269.01		

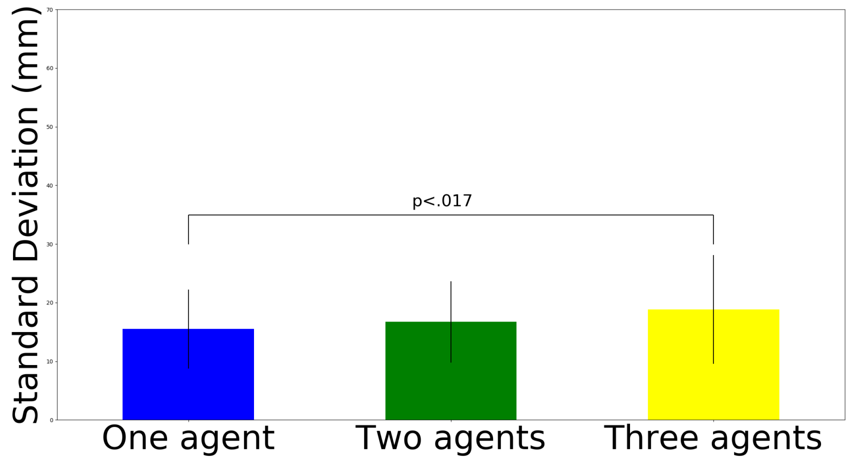

Figure 3.10: Standard deviation results on orthogonal horizontal movement interference.

A three-way ANOVA for vertical movement in movement interference was conducted, of which the results are shown in Table 3.2. There was not significant interaction effect found on the variables. On the other hand, a main effect (F (2, 276) = 4.49, p<.05) in the number of robots was found. Then, a post-hoc analysis utilizing Bonferroni correction (here, adjusted alpha levels of 0.017) was conducted (Figure 3.11). The results show that three agent set-up has greater interference in the movement (M=18.82 mm, SD=9.28 mm) than the one agent set-up (M=15.49 mm, SD=6.76 mm), where t (190) =2.85, p=.005, Cohen's d=0.41. In contrast, the two agent set-up (M=16.72 mm, SD=6.93 mm) was not sig-nificantly different from the one agent set-up (M=15.49 mm, SD=6.76 mm), where t (190) =0.18, p=.215, Cohen's d=0.18; similarly, three agent set-up (M=18.82 mm, SD=9.28 mm) was not significant different from the two agent set-up (M=16.72 mm, SD=6.93 mm), where t (190) =1.78, p=.076, Cohen's d=0.26.

Table 3.2: Three-way ANOVA for vertical interference movement.

	df	sum_sq	mean_sq	F	PR(>F)
C(Number of Agents)	2	546.08	273.04	4.49	0.01
C(Type of avatar)	1	0.01	0.01	0	0.99
C(Type of movement)	1	3.75	3.75	0.06	0.8
C(Number of Agents):C(Type of avatar)	2	153.58	76.79	1.26	0.28
C(Number of Agents):C(Type of movement)	2	15.36	7.68	0.13	0.88
C(Type of avatar):C(Type of movement)	1	74.33	74.33	1.22	0.27
C(Number of Agents):C(Type of avatar):C(Type of movement)	2	35.41	17.7	0.29	0.75
Residual	276	16790.12	60.83		

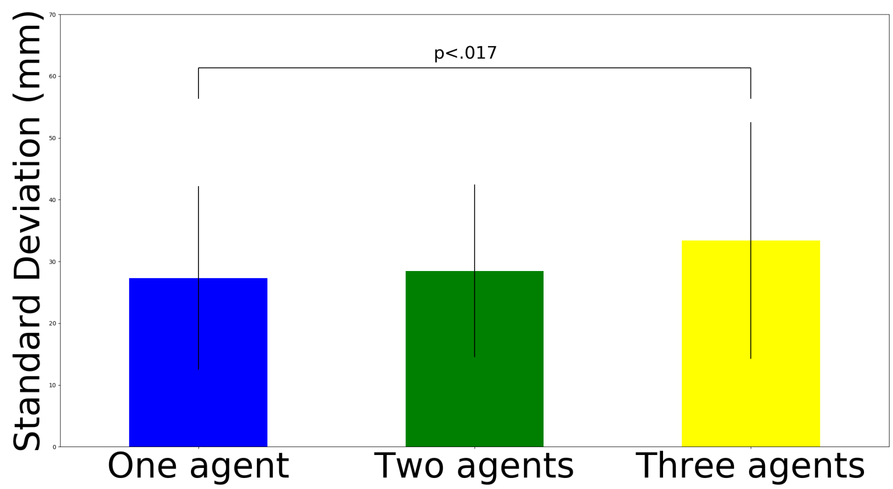

Figure 3.11: Standard deviation results on orthogonal vertical movement interference.

3.2.2 Results on time Synchronization

A three-way ANOVA for the vertical time synchronization was conducted, the results of which are shown in Table 3.3. There was not significant interaction effect found on the variables. On the other hand, a main effect (F (1, 276) = 10.63, p<.01) in the agents' avatar was found. Then, post-hoc analysis utilizing Bonferroni correction (here, adjusted alpha levels of 0.05) was conducted. (Figure 3.12). The results show that robot has greater phase change in the hand movement (M=44.66 deg, SD=33.24 deg) than the human agent set-up (M=31.68 deg, SD=33.64 deg), where t (286) =3.29, p<0.01; Cohen's d=0.39. In contrast, there was no sig-nificant main effect for the agent number (F (2,276) = 0.84, p=0.43); similarly, there was no significant main effect for the biological movement (F (1,276) = 1.06, p=0.3).

Table 3.3: Three-way ANOVA for vertical on time Synchronization.

	df	sum_sq	mean_sq	F	PR(>F)
C(Number of Agents)	2	1274.51	637.26	0.84	0.43
C(Type of avatar)	1	10190.42	10190.42	13.39	0
C(Type of movement)	1	803.9	803.9	1.06	0.3
C(Number of Agents):C(Type of avatar)	2	383.85	191.93	0.25	0.78
C(Number of Agents):C(Type of movement)	2	641.52	320.76	0.42	0.66
C(Type of avatar):C(Type of movement)	1	170.62	170.62	0.22	0.64
C(Number of Agents):C(Type of avatar):C(Type of movement)	2	812.26	406.13	0.53	0.59
Residual	276	210073	761.13		

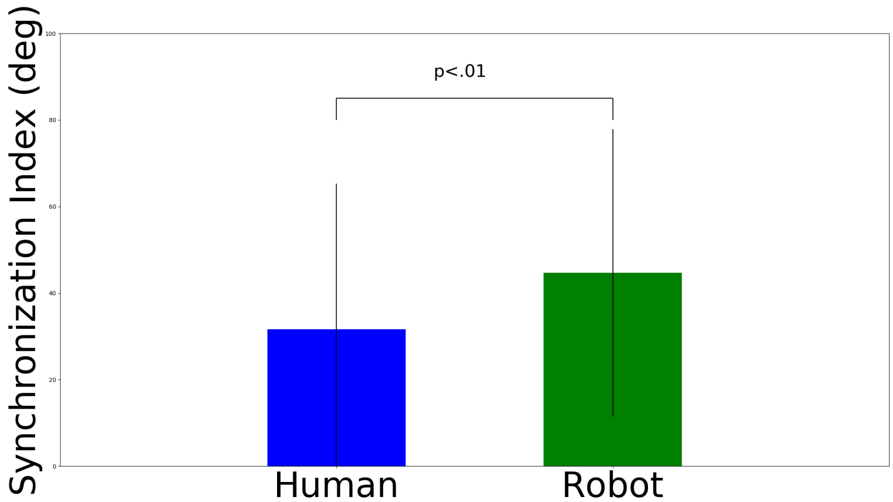

Figure 3.12: Synchronization index results on orthogonal vertical movement interference.

A three-way ANOVA for the Horizontal time synchronization was conducted, the results of which are shown in Table 3.4. There was not significant interaction effect found on the variables. On the other hand, a main effect (F (1, 276) = 13.39, p<.01) in the type of avatar was found. Then, post-hoc analysis utilizing Bonferroni correction (here, adjusted alpha levels of 0.05) was conducted. (Figure 3.13). The results show that robot has greater phase change in the hand movement (M=34.72 deg, SD=30.99 deg) than the human agent set-up (M=22.82 deg, SD=23.18 deg), where t (286) =3.69, p<0.01; Cohen's d=0.39. In contrast, there was no significant main effect for the agent number (F (2,276) = 0.12, p=0.89); similarly, there was no significant main effect for the biological movement (F (1,276) = 0.35, p=0.55).

Table 3.4: Three-way ANOVA for Horizontal interference movement.

	df	sum_sq	mean_sq	F	PR(>F)
C(Number of Agents)	2	274.95	137.48	0.12	0.89
C(Type of avatar)	1	12139.16	12139.16	10.63	0
C(Type of movement)	1	401.44	401.44	0.35	0.55
C(Number of Agents):C(Type of avatar)	2	771.85	385.92	0.34	0.71
C(Number of Agents):C(Type of movement)	2	25.15	12.58	0.01	0.99
C(Type of avatar):C(Type of movement)	1	2759.06	2759.06	2.42	0.12
C(Number of Agents):C(Type of avatar):C(Type of movement)	2	469.99	234.99	0.21	0.81
Residual	276	315197.4	1142.02		

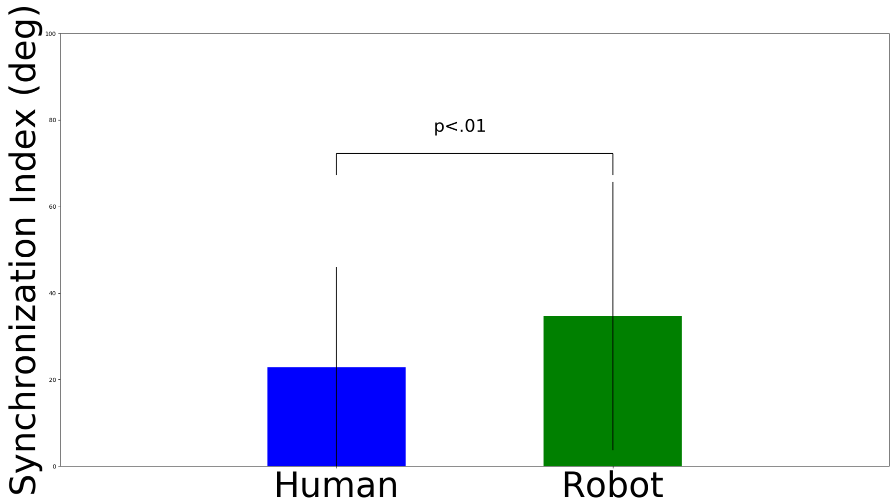

Figure 3.13: Synchronization index results on orthogonal horizontal movement interference.

3.3 Discussion

Results of the experiments revealed that the effective factors for movement interference and synchronization index are different. It was found that H1 was partially supported while H2 and H3 were not supported. On the other hand, H5 was supported while H4 and H6 were not supported. In the following subsections, the possible reasons for the difference are discussed based on each factor.

3.3.1 Number of agent

As it is known, the ability of a group of people to accomplish a task efficiently is known to be enhanced when they have a similar efficiency level [117, 118]. Additionally, members of groups with different beliefs are more likely to change their beliefs when the rest of the group gradually changes their own [119]. Moreover, emotions are easily spread within a group when there are more people [120]. Thus, by increasing the number of agents, the participants might perceive themselves as part of the group of agents, effectively affecting their hand movement. However, there was no evidence that the two-agent set-up increased interference compared to the one-agent set-up nor that the three-agent set-up increased interference compared to the two-agent set-up. The reason for not having a significant difference might be because in a two-agent set-up the number of the agents were not conceiving a group. It is possible that in the two-agent set-up the two agents were perceived as two individuals instead of a single group due to the small number of members. Therefore, it is suggested that for recognizing a group, a minimum of three agents are required for affecting the movement interference. Nevertheless, the effect of the number of agents was not observed for the synchronization. In the interference experiment, the movement of the subject and agent(s) were orthogonal while for the synchronization experiment, it was parallel. When subjects were looking at orthogonal movements, they may focus less on the rhythm (since it appears to be difficult to follow two or more orthogonal movements by eye) and place a larger emphasis on the properties of the projected agent (such as the number of projected agents). In contrast, in the case of the parallel movement, the subjects might focus on the rhythm of the movements (since focusing on and following the parallel movements by eyes seems not to be so difficult) but not on the features of the projected agents. As a result, for the interference experiment, the group of the agents might be recognized easily by the participants while for the synchronization experiment it could be difficult. This could be the main reason that the factor of the number of agents worked for the interference but has not appeared in the synchronization experiment.

3.3.2 Type of Avatar

The effect of the type of avatar may be due to the perception of the subject about the agent as a peer when it has similar physical aspect doing the same action. As it is known, people feel more motivated to use avatars when the avatar share similar characteris with them [121] and humans are more willing to cooperate by synchronizing their starting-ending actions on cooperative spaces when they share similar characteristics [122]. Therefore, in the synchronization experiment, more similar agent, i.e. the human agent, led to more synchronized movement by the participant.

3.3.3 Biological movement

The effect of the biological movement was not observed in any of the experiments. Even though Kilner[28] proposed that biological movements might have a significant effect on the interference of the moving hand, in our case it seems that too many conditions were combined. Therefore, the difference in movement seems less obvious to the participant. As a consequence, neither the movement interference nor the antiphase synchronization was affected by the agent's biological movement.

3.3.4 Synchronization Index

Previous studies verified the effect of biological movement between robots and humans [28, 62, 63], the robotic shape influence [107], and humanoid robots influence [117]. However, they did not study how the agents influence the synchronization of the action. This synchronization influence needs more research to clarify the details of the nature of its influence. In this chapter, the proposed index is a phase-based index that was designed for this experiment. To study more deeply the synchronization concept, more studies with different measurements might be beneficial for understanding the dimensions of the influence of the agents on this field.

3.4 Limitations

Although in this experiment the influence of the number of agents and the type of movements were explored, it was not tested any social behavior of the robots such as speech, gaze, behavior, or relational interaction among the agents. Subjects might have been affected more significantly if such social behaviors had been included. Moreover, fatigue may influence the results of movement interference. In addition, even though it was experimented using counterbalance, the subjects participated in all the set-ups, being therefore possibly influenced by previous conditions.

3.5 Conclusions

This chapter, it was proposed that a group of projected agents moving their hands orthogonally to the subject would be able to increase movement interference. The projected agents were: one, two, or three agents; humans or robots; biological agent movements or linear agent movements. To explore the effect of the proposed factors, the following comparison was made: 1) the effect of exposing subjects to a different number of agents, different type of avatar, and different types of movements in incongruent vertical or horizontal subject movement. 2) the effect of exposing subjects to a different number of agents, different type of avatar, and different types of movements in congruent vertical or horizontal subject movements. In these experiments, it was gathered 24 participants that were involved on both experiments that last around 25 minutes each. The interference of the agent on the hand movement of the participants was studied which was around 13 mm. Also, the synchronization of the agent on the hand movement of the participant was studied which was around 12 deg. Increased movement interference was found between the three robots' set-up and the one robot set-up. It was also found, that enhanced synchronization between human set-up and robot set-up. The results showed that the numbers of robots enhanced the interference movement which reveals the knowledge about the influence that number of agents in the interference of the movement of people. Moreover, it was proposed an additional paradigm that was not previously considered which was the influence on the synchronization in which it was found that the avatar type influence on keeping the synchroni-zation. The results reveal the knowledge about the influence that the avatar might have on influencing people for performing complex tasks such keep synchronization. However, this experiment did not verify the effect of different social behaviors of agents or different types of non-humanoid avatars, which will be analyzed in forthcoming experiments. It is also worth to evaluate the group dynamics on the interference of synchronization and movement on people. The combinations of different type of non-humanoid avatars would help to understand these group dynamics on depth.

CHAPTER 4

STUDY II:INFLUENCE OF AGENTS IN SYNCHRONIZED MOVEMENTS WITH MUSICAL BACKGROUND

In this chapter, it was conducted an experiment to evaluate the effect of the number of robots in synchronizing rhythmic humans movement. Two persons participated in the experiment: one experimenter and one subject. They were positioned in front of each other and their task was to play drums in a musical environment; the robots were moved according to the drumbeats. The level of synchronization among the humans and the level of enjoyment experienced was measured. To measure the level of synchronization, it was calculated the variance of the timing difference between the first drumbeat of the experimenter and the next nearest drumbeat of the subject and multiplied the variance result by minus one.

4.1 Experiment 1

4.1.1 Method

To investigate the effect of multiple robots on human synchronization, it was conducted an experiment in which three conditions were adopted: zero robots, one robot, and three robots. As argued for in the Introduction, it is hypothesized that (H1) the higher the number of robots, the higher the level of synchronization will be between the two persons. Meanwhile, a positive influence is considered to occur in human perception of human–robot interaction with the increase in synchronization. In other words, it is hypothesized that (H2) the higher the number of robots, the greater the enjoyment perceived by the subject. The subjects provided written consent, and the experiments were approved by the ethical committee of the Graduate School of Engineering Science, Osaka University.

4.1.2 Subjects

Twenty-six people participated in all the conditions for the first experiment (13 females and 13 males, mean age = 22.6, SD = 3.9). One female and one male subject were excluded from the analysis because they could not complete the experiment owing to technical or environmental problems. The order of the conditions assigned to the subjects was randomly selected so that it was counterbalanced.

4.1.3 Experimental Setup

It was developed a musical system using humanoid robots that could produce rhythmic body movements according to signals detected from electric drums, which were played by the experimenter and the subject. Figure 4.1 illustrates a sample scene of the zero-, one- and three-robot conditions. One and three robots were placed on a table between the experimenter and the subject in the one- and three-robot conditions, respectively. In the one-robot condition, two robots were moved under the table. In the zero-robot condition, one robot was left on the table but it was covered by a small panel and the height of the table was reduced to conceal this robot from the subject's view.

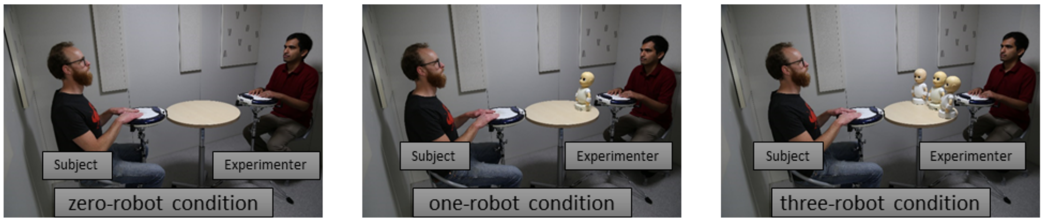

Figure 4.1: Sample scene of the zero-, one-, and three-robot conditions.

Humanoid Robots

For the one- and three-robot conditions, it was used a robot, "CommU," developed in collaboration with Osaka University and Vstone Co., Ltd. CommU is a tabletop humanoid robot (see Figure 4.2) with a height of 304 mm. Its head employed 3 degrees of freedom (DoF), two eyes with 3 DoF, upper eyelids with 1 DoF, a mouth with 1 DoF, two arms, each with 2 DOF, and a waist with 2 DOF. CommU is programmable with two types of movements to represent rhythmic behavior: neck and arm movements. Neck movements were produced with the roll joint of the neck, swinging it left and right with an amplitude of 20°. Arm movements were produced with pitch joints, alternately swinging them up and down with an amplitude of 30°. These robot movements can be triggered independently by external signals.

Figure 4.2: CommU.

Rhythm Detection System

The rhythm detection system detected the onsets of drumbeats played in real-time on two electric drums (KORG WAVE DRUM WD-X, 75 mm (H) x 349 mm (W) x 344 mm (D)). The drum signals were captured by an external sound card (Presounous 44VSL) and delivered to program units using a sound server daemon, "Jack server audio." It was used to detect the onset of drumbeats in audio windows signals of 22 ms. The onset of each sound was calculated based on the energy contour of the waveform to detect percussive sounds [123], which first calculates the fast Fourier transform of the sound signal within the focused time window. It later calculates the integral of the amplitude along the frequency axis. Finally, it identifies the onset if it surpasses a predefined threshold. The identified onset of drumbeats was used to produce robot movements: one on the experimenter's drum (neck movements) and one of the subject's drum (arm movements). It was selected these two movements because they are considered to be simple enough to show the rhythms of the subjects and the experimenter and could be produced without any interference between them. In other words, the robot exhibited a combination of both the rhythms of the subject and the experimenter such that it seemed to produce the neck movements in the same rhythm as the experimenter as well as the arm movements in the same rhythm as the subject. It was adopted the KXstudio Linux operating system, which uses a special kernel to achieve a faster audio processing.

Finally, in order to control the potential effect of the experimenter's gaze among subjects, it was placed a light-emitting-diode (LED) on CommU to light up at preprogrammed intervals to let the experimenter know the timing for changing his gaze. When the LED was on, the experimenter had to watch the subject; when it was off, he had to watch the table. The LED was placed such that the subject could not directly see it. The LED was alternately turned on and off every 20 s.

4.1.4 Environment

The experimental space was built in a sound-proof room (AMCVB37, YAMAHA Co.). It spanned 2.1 m (W) x 2.6 m (D) x 2.1 m (H). A heightadjustable table was used for placing the robot. Chairs were installed for the subject and the experimenter, each in front on the opposite sides of the table. The drums were placed 20 cm in front of each person. To record the experiment, it was mounted two cameras to capture the drumbeats, and it was recorded the drum sounds. A personal computer was placed in the room so that the subject could search and listen to music, choosing a favorite song to be used in the experiment. The subject stood 90 cm down from the table, whereas the experimenter stood 60 cm from the table.

4.1.5 Procedure

All the instructions, including informed consent, were provided via written documentation. The subjects' tasks included attending musical sessions in which they played a drum with an experimenter who played on the other side of a table on which zero-to-three robots stood. In the one- and three- robot conditions, the robots were placed on the table so that the subjects faced them and could clearly see them. The subjects' tasks also included evaluating the impressions formed during each session. The experimenter demonstrated a typical method of playing drums as training. The experimenter needed to produce a rhythm that would not be spontaneously produced by subjects. In this experiment, the experimenter was trained to tap twice at each beat of the song. Figure 4.3 shows an example of how the experimenter and a subject played. Approximately, the experimenter (green circles) tapped twice every 1.2 s while the subject (blue squares) first tapped every 1.1 s, but later changed to every 0.6 s.

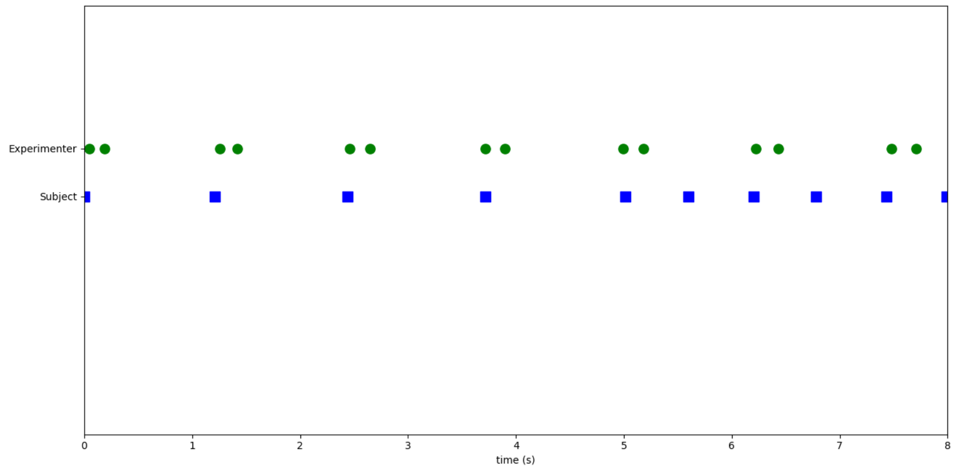

Figure 4.3: Examples of the tapping rhythms of the experimenter (green circles) and a subject (blue squares).

The subjects were asked to choose their favorite songs to reduce fatigue due to the long duration of the experiment in which they had to listen to the same song. The experimenter edited the sound files so that they lasted only 3 min each. The subjects were then asked to play a drum by following the rhythm of their selected song in their preferred timing while the experimenter tapped his own at the predefined timing. Then, to reduce possible novelty effects, it was asked the subjects to attend four 3 min sessions and they were habituated to the situations where they played drums with the experimenter and with different numbers of robots. The subjects attended a training session with two robots, then they attended three sessions to habituate to the novel environment they would be exposed to during the experiments. During the experiments, they played drums with zero, one, or three robots. The order for the three training sessions was random as were the subsequent experiment sessions. In the latter sessions, the order of the number of robots was counterbalanced. Upon completing each session, each subject had 20 s to answer a question on enjoyment. The subjects then had a 15 s rest. When the habituation process has been completed, each subject was allowed to rest for 2 min. In total, experiment 1 lasted 40 min. The procedure for experiment 1 was the same as that of the counterbalanced habituation.

4.1.6 Evaluation

One question to evaluate the subject's perception of the session was asked, and it was processed the timing of drumbeats to evaluate the level of synchronization. The question was: "Please score how much you enjoyed each session by choosing one number from one to five where '1' indicates that 'you did not enjoy it at all,' '3' is 'undecided,' and '5' is 'you enjoyed it very much.'" It was detected the beat rhythm of the experimenter and the subject where the detected beats were the starting time of the movements of the robot after they tapped the drums. The sequences of the detected beats were preprocessed to interpret the turn-taking between the experimenter and subject. It was filtered out every beat except for the first beat on the experimenter playing cycle (experimenter's i-th double tapping.).

The experimenter was trained to tap a drum two times for each beat in the selected song. Hence, the first tapping of the experimenter for every beat and the tapping of the subject appeared just after they were extracted and paired. Then, it was analyzed the sequence of the time difference between these paired beats. Hereafter, the beats of the experimenter and the subject in each pair are referred to as the leading beat and the turn-taking beat, respectively. Then, it was analyzed the sequence of the time difference between the filtered beats and the subject's turn-taking beat as equation 4.1 shows.

$$D_i = S'_i - E'_i \tag{4.1}$$

where E'_i is the first timing of the experimenter's i-th double-tapping, S'_i is the first timing of the subject's taping after the first timming of the experimenter's i-th double-tapping E'_i, and D_i is the discrepancy for the i-th tapping.

An example of how the E'_i is obtained is shown on figure 4.4. In a similar way, an example of how S'_i is obtained based on E'_i is shown on figure 4.5. Finally, the equation 4.1 is shown on figure 4.6.

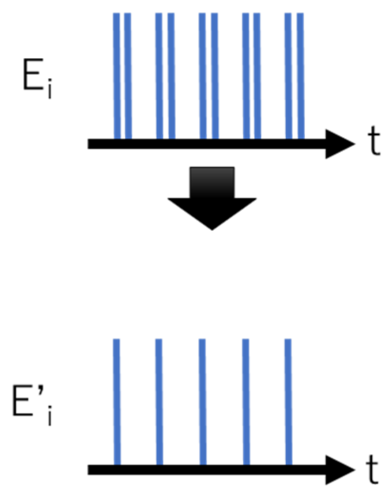

Figure 4.4: Example of the conversion of the first timing of the experimenter's i-th double-tapping

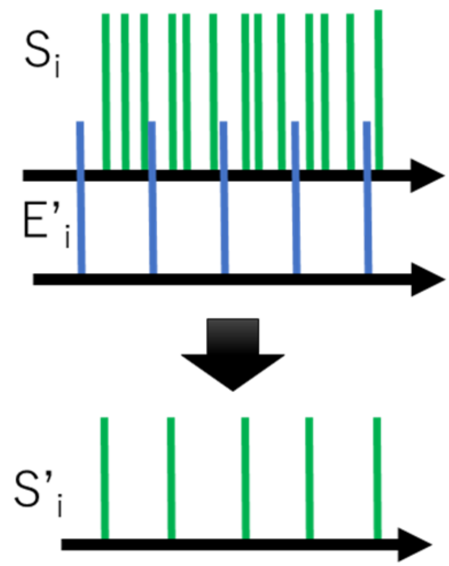

Figure 4.5: Example of the conversion of the first timing of the subject's taping after the first timming of the experimenter's i-th double-tapping

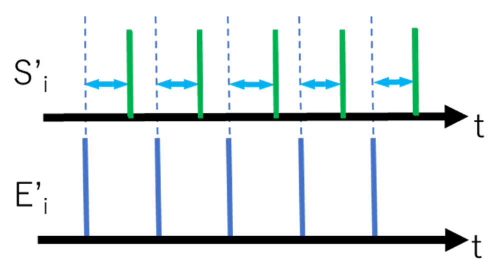

Figure 4.6: Visual example of the discrepancy for the i-th tapping.

Finally, it was calculated the variance and multiplied it by minus one because less variance in the time difference seems to represent reliability of synchronization as the equation 4.2 shows. Because synchronization between two persons takes time, it was used the last 48 s of each session for analysis.

$$Sl = -\frac{\sum_{i=1}^{N}(D_i - \overline{D})^2}{N} \tag{4.2}$$

where Sl is the synchronization level, n is the time discrepancy amount of data and D the discrepancy discrete signal and $\overline{D}$ is the average of D.

4.1.7 Results

The time sequences of the variance of the time difference in the robots' conditions are plotted in Figure 4.7. The average value for the three-robot condition was significantly greater than that of the zero-robot condition (t(23)=2.64, p<.05) after Bonferroni correction. On the other hand, there were no significant differences between the zero- and one-robot conditions ((t(23)=0.58, p=.566) or the one- and three-robot conditions (t(23)=1.31, p=.204). The enjoyment of the zero-robot condition was significantly smaller than that of the one-robot condition (t(23)=3.50, p<.05) after Bonferroni correction. It was also significantly smaller than that of the three-robot condition (t(23)=3.84, p<.01) after Bonferroni correction. The enjoyment of the one-robot condition was not significantly smaller than that of the three-robot condition. Figure 4.8 plots the enjoyment results for the number of robot conditions.

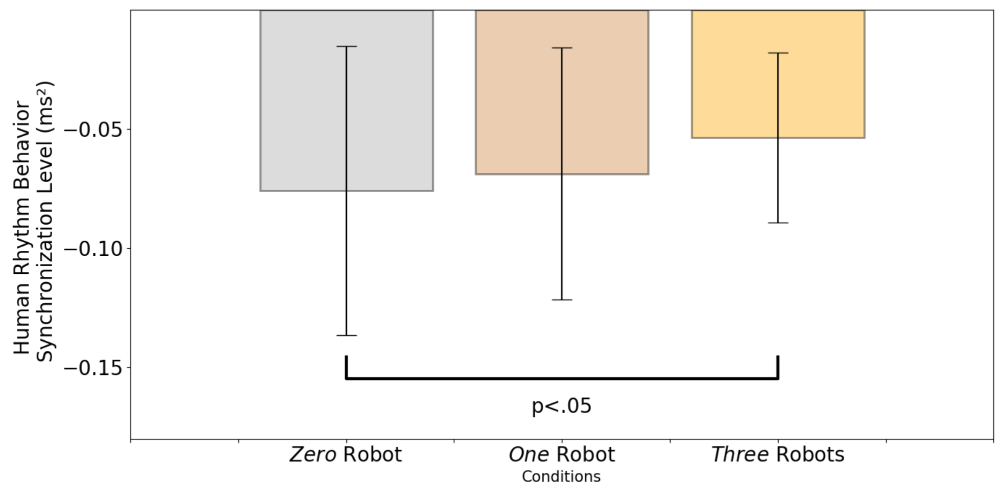

Figure 4.7: Human–Human Synchronization Level according to the Number-of-Robot Conditions.

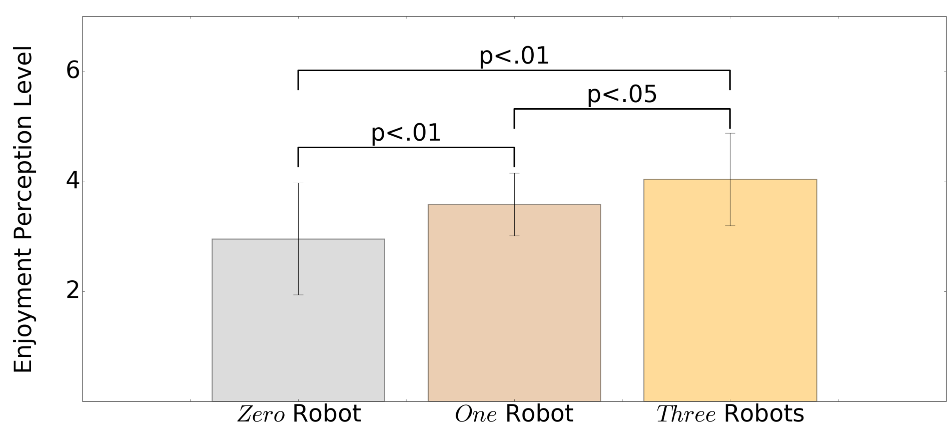

Figure 4.8: Enjoyment Perception according to the Number-of-Robot Conditions.

The Pearson correlations between the levels of synchronization and enjoyment for all conditions were not significant: 0.22 (n.s.) in the zero-robot condition, -0.22 (n.s.) in the one-robot condition, and 0.19 (n.s.) in the three-robot condition.

4.2 Discussion

The levels of synchronization reported in the experiment suggest that three robots influenced human synchronization more compared with no robot. However, no significant differences were observed in other pairs. In other words, hypothesis H1 is partially supported. Moreover, it was found that enjoyment perception of the interaction was enhanced when more robots were added; meaning that H2 was supported. It is known that mood contagion in groups is higher in collaboration tasks when the group has more than two individuals [120]. The more the people in a group look toward a certain direction, the more the audience will also mimic them by looking in the same direction [124]. Thus, a group of robots may have the same influence on humans and prompt a person to mimic their behavior. In this chapter, it was presumed that the subjects, robots, and experimenter all considered themselves to be part of a group of collaborative musical session. Therefore, they were prompted to mimic the group rhythm produced by the experimenter, and this was more pronounced in the three-robot condition. Individuals who are not identified as part of a group have less social influence than those identified as part of a group [41]. The lack of significant differences between the zero-condition and the one-robot condition may imply that subjects in the one-robot condition did not perceive the robot or the experimenter as part of one group. Meanwhile, it is possible that some subjects had such a perception of the group not only in the three-condition but also in the one-robot condition, which can possibly lead to lack of significant difference between the one-condition and the three-robot condition.

4.3 Limitations

In the discussion of Experiment 1, the results suggest that the actual number of robots, that is, zero, one, or three, may influence its perceived sociability. Hence, the number of robots may influence the synchronization of two people participating in the experiment. Other studies suggest that nonsocial individual stimuli such as a bouncing ball, or the sound of a metronome can influence synchronization between people, including their synchronization to the beat of a song [125, 126]. However, in the present experiment, it was not evaluated the case where the robot influenced the synchronization of subjects with the experimenter only due to the movement.

4.4 Conclusions

In this chapter, it was proposed that multi-party interaction that includes people and robots can increase the level of synchronization in a musical performance. It was developed a multi-robot system in which humanoid robots moved according to the drums played by two people. Using multiple robots the synchronization is enhanced.

CHAPTER 5

STUDY III: INFLUENCE OF AGENTS IN SHARING SYNCHRONIZED MOVEMENTS IN A ROBOTIC BODY

In this chapter, it was evaluated the effect of robots using their bodies to synchronize with humans. To support the interpretation of the evaluation of this experiment in relation to the rubber hand illusion, it was conducted a second experiment to evaluate the effect of an in-sync robot that moves in-sync with the subject's rhythm by comparing it with an out-of-sync robot that moves with a fixed rhythm independent of the subject. Two persons participated in the first experiment: one experimenter and one subject. They were positioned in front of each other and their task was to play drums in a musical environment; the robots were moved according to the drumbeats. It was measured the level of synchronization among the humans and the level of enjoyment experienced. The level of synchronization among the humans and the level of enjoyment experienced was measured. To measure the level of synchronization, it was calculated the variance of the timing difference between the first drumbeat of the experimenter and the next nearest drumbeat of the subject and multiplied the variance result by minus one. In the last experiment, only one subject and one robot were involved. The subject tapped a drum in front of the robot, which at first moved rhythmically but suddenly stopped after a while. It was measured how the average time lag between the subject's tapping increased between the periods before and after the robot stopped.

5.1 Experiment 1

5.1.1 Method

To investigate the effect of the type of rhythm on the synchronization of humans, it was conducted an experiment where two humans and three robots moved. Two conditions were adopted. In the mixed-rhythm robot condition, robots move their bodies using both the experimenter's and the subject's rhythms, whereas in the unmixed-rhythm robot condition, they move using only the experimenter's rhythm. It was did not consider the condition where robots used the rhythm of only the subject. This was because the main interest of the experiment was the influence of the experimenter's rhythm on the subject, depending on whether the robot's representation included the subject's rhythm. As discussed in the introduction, if the subject observes part of the robot movement synchronizing with subject's rhythm, the subject would be more influenced by the rest of the movement. In other words, it is hypothesized that (H1) the level of synchronization between the experimenter and the subject is greater in the mixed-rhythm robot condition than in the unmixed-rhythm condition. Meanwhile, as in Experiment 1, a positive influence is considered to occur in human perception of human–robot interaction with the increase in synchronization. Therefore, it is also hypothesized that (H2) the subjects feel greater enjoyment in the mixed-rhythm condition than the unmixed-rhythm condition.

5.1.2 Subjects

Twenty-six persons (13 females and males: mean age = 22.6, SD = 3.9) participated in both sessions of the mixed-rhythm and unmixed-rhythm conditions of the second experiment. The subjects who participated in the experiment of the previous chapter were also involved in this experiment. The order of the conditions assigned to the subjects was randomly selected so that it was counterbalanced.

5.1.3 Experimental Setup

In the mixed-rhythm condition, the robots exhibited a combination of both the rhythms of the subject and the experimenter, just as in Experiment 1. That is, their neck moved in sync with the experimenter's drumbeats while their arms moved in sync with those of the subject. On the other hand, in the unmixed-rhythm condition, the robots only exhibited the rhythm of the experimenter. That is, both their necks and arms moved in sync with only the experimenter's drumbeat.

5.1.4 Procedure

All the subjects participated in the experiment of the first chapter participated on this experiment. The experiment was conducted after a 2 min break. Each subject was exposed to a second habituation process of four sessions. The first two sessions were performed using only one hand under the conditions of mixed and unmixed rhythms. The other two sessions were performed using two hands in the same order as the previous two. When the habituation process has been completed, the subjects were allowed to rest for 2 min. In total, this experiment lasted 35 min. The experiment had the same order as the habituation process, which was counterbalanced.

5.1.5 Evaluation

The same question as in the previous chapter was asked to evaluate the subject's perception during sessions. The same analysis as in study 2 was performed to evaluate the level of synchronization (Equation 4.2).

5.1.6 Results

The time sequences of the variance of the time difference in the mixed and unmixed rhythm conditions are plotted in Figure 5.1. In the last 48 s of the session, synchronization between the subjects and the experimenter in the unmixed rhythm condition (M=0.06, SD=0.05) and mixed-robot condition (M=0.07, SD=0.08) resulted in different trends in the expected direction without reaching significance; $t(23)=1.83$, $p < 0.1$. The enjoyment from the mixed-rhythm condition (M=4.0, SD=0.96) was significantly larger than that from the unmixed rhythm condition (M=3.16, SD=1.4) ($t(23)=3.5$, $p<.01$), as plotted in Figure 5.2.

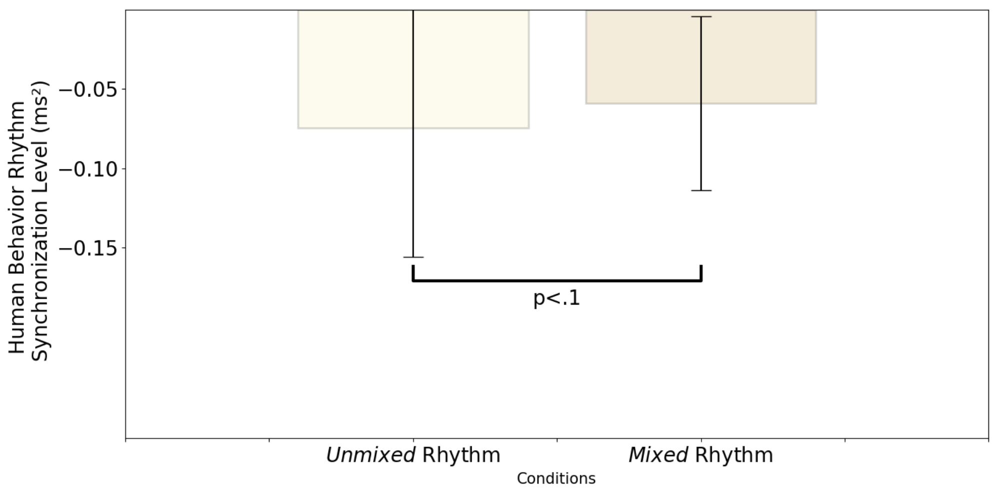

Figure 5.1: Human–Human Synchronization Level according to the Rhythm Condition.

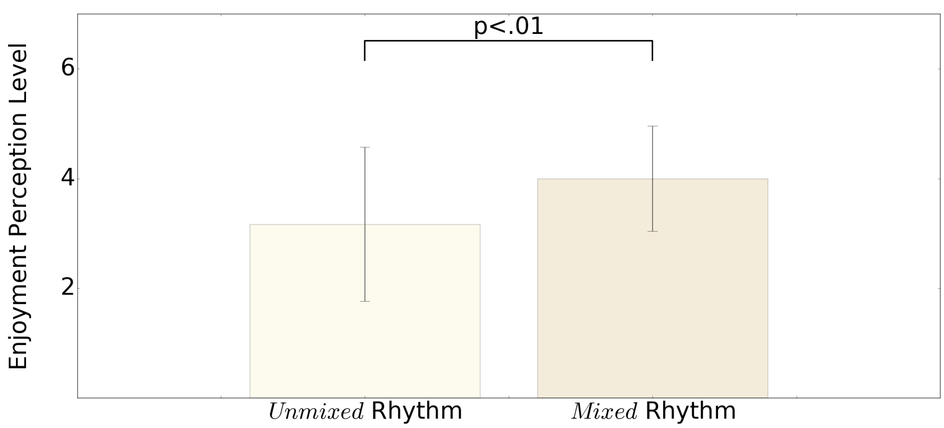

Figure 5.2: Enjoyment Perception during the Rhythm Conditions.

The Pearson correlations between the levels of synchronization and enjoyment in both conditions were insignificant: -0.34 (n.s.) in the mixed-rhythm condition and 0.19 (n.s.) in the unmixed-rhythm condition.

Session type	Experiment 1	Experiment 2
Habituation session	Two robots	Two robots mixed rhythm
	Condition A	Two robots unmixed rhythm
	Condition B	Condition D
	Condition C	Condition E

Experimental session	Condition A	Condition D
	Condition B	Condition E
	Condition C	

*Conditions A, B, C could be Zero-, One-, or Three-robot condition.
Conditions D, E could be Three robots mixed or unmixed rhythm condition.

Figure 5.3: Experiment 1 and Experiment 2 order sequence.

5.2 Experiment 2

5.2.1 Method

It was investigated whether a robot that has moved in-sync with the drumbeats of a subject becomes effective in influencing them, which is considered to be similar to the phenomenon of the rubber hand illusion. It was conducted an experiment where a subject played a drum in front of a robot. It was adopted two conditions: the in-sync robot condition where the robot moved to follow the subject's rhythm and the out-of-sync robot condition where it moved to follow a fixed rhythm independent of the subject. If a similar phenomenon as the rubber hand illusion is induced, it would be expected that the sudden changes occurring to the robot's rhythm would be confused with the subject's rhythm and be followed by the subject. In other words, it is hypothesized that (H3) the subject in the in-sync robot condition changes subject's rhythm more than that in the out-of-sync condition. Meanwhile, as in previous experiments, a positive influence is considered to occur in human perception of human–robot interaction with the increase in synchronization. Therefore, it is also hypothesized that (H4) the subject will feel greater enjoyment in the in-sync robot condition than in the out-of-sync condition.

5.2.2 Subjects

Twenty-two persons (7 females and 15 males: mean age = 25.58, SD = 3.26) participated in the sessions, alternately repeating in-sync and out-of-sync robot conditions two times. The subjects were not involved in the previous two experiments. The order of the conditions assigned to the subjects was randomly selected so that it was counterbalanced. The experiment was approved by the ethical committee of the Graduate School of Engineering Science, Osaka University.

5.2.3 Experimental Setup

The subject was asked to tap a drum with only one robot in the room. The experiment ran over a short time span and a song called "Macarena" was repeatedly played throughout the experiment in an attempt to keep the subject alert. Unlike in Experiments 1 and 2, there was no need for the subject to choose a favorite song because of the duration of the experiment. In the in-sync robot condition, it moved its neck and hands each time the subject tapped the drum. In the out-of-sync robot condition, the robot was programmed to move in a specific rhythm and the subjects were not able to move the robot during the session. The rhythm selected was similar to the one provided by the experimenter in the first and second experiments.

5.2.4 Environment

The experimental space was built in a space of 2.0 m (W) x 4.0 m (D) x 3.4 m (H). The drum and the robot were placed on a simple table. Chairs were placed in front of the table for the subject. The drums were placed 20 cm in front of the subject, and the robot was 10 cm behind the drums. To record the experiment, it was mounted one camera to capture the drumbeats.

5.2.5 Procedure

The subject was asked to read an instruction document about the experiment. In the document, the subjects were asked to play a drum until the song ends. The song lasts for 60 s; the robot moves only in the first 50 s, while it suddenly stops in the remaining 10 s. The subjects attended four sessions in which the first two sessions were the habituation sessions and the last two were the experimental sessions. Both the habituation and experimental sessions consist of in-sync and out-of-sync conditions. After each session, the subject was asked a question about the enjoyment he or she experienced. Figure 5.4 shows a sample scene of how the subject played. The in-sync and out-of-sync conditions were alternately conducted, and the order of the condition was counterbalanced.

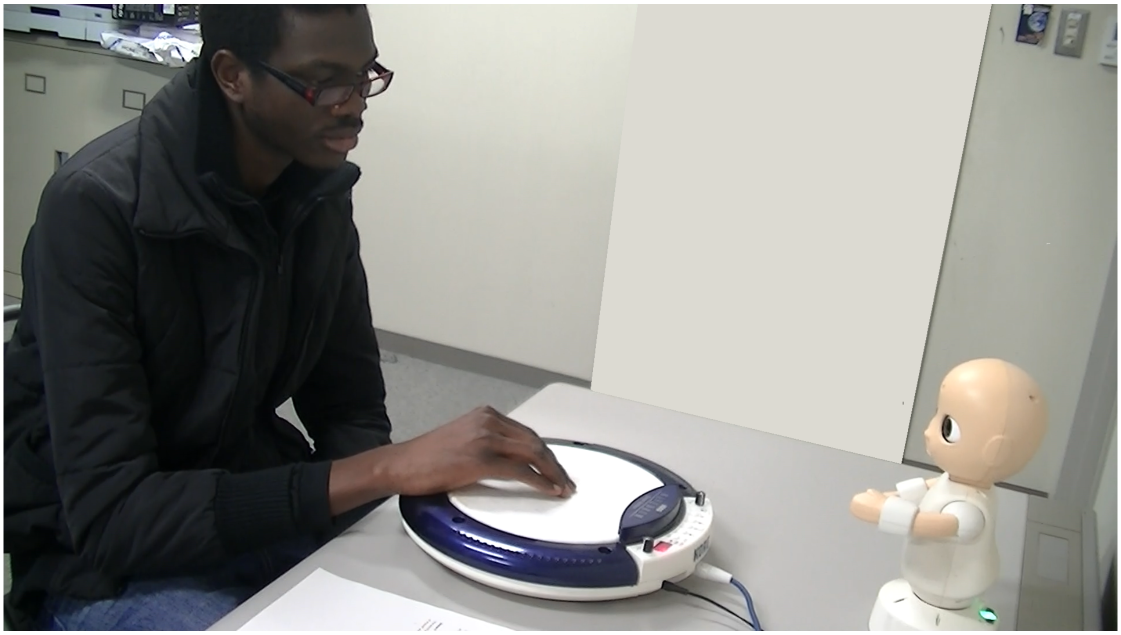

Figure 5.4: Sample scene of a subject tapping in front of a robot.

5.2.6 Evaluation

It was used the same question used in the previous experiment to evaluate the perception of enjoyment during each session. In addition, it was calculated how the average time lag between the subject's tapping increased from the former session in which the robot was moving to the latter session after it stopped. That is, it was calculated by dividing the average time lag between the tapping in the first 50 s session by that in the last 10 s session as the equation 5.1 shows. It was compared the ratios using the paired t-test.

$$Sl = \frac{(t_N - t_n) \times (n-1)}{(t_n - t_1) \times (N-n)} \tag{5.1}$$

where Sl is the synchronization level, n is amount of tapping performed on the first 50 seconds, N is the total amount of tapping, and t is the tapping time.

5.2.7 Results

The increase in the ratio of the average time lag in both conditions are shown in Figure 5.5. It is larger in the in-sync robot condition (M=1.06, SD=0.16) than in the out-of-sync robot condition (M=0.97, SD=0.11) (t(21)=2.08, p<.05). The enjoyment of the robot moved by the subject's rhythm (M=3.73, SD=1.03) was significantly larger than that of the robot moved by the predefined rhythm (M=3.05, SD=1.05) (t(21)=2.73, p<.05). This result is plotted in Figure 5.6.

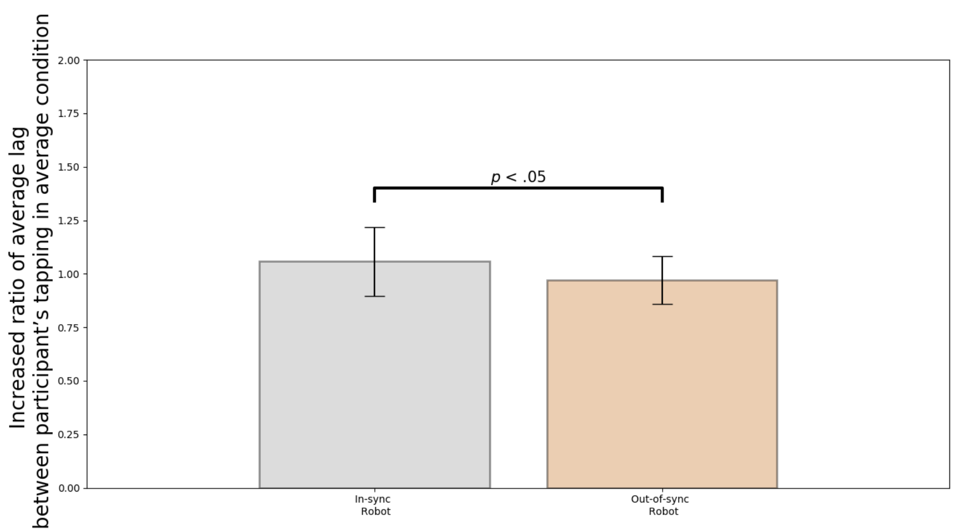

Figure 5.5: Increase in the ratio of the average time lag between subject's tapping average condition.

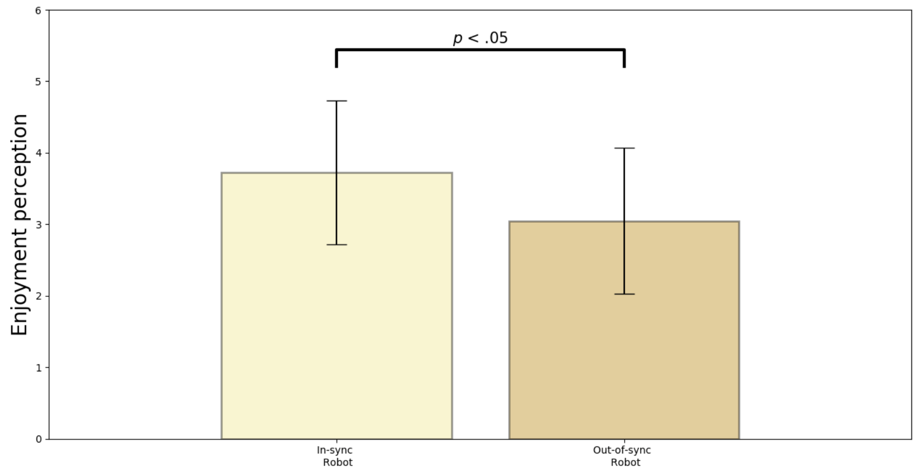

Figure 5.6: Enjoyment perception in experiment 3.

The Pearson correlations between the increase in the ratios of the average time lag between tapping and enjoyment were not significant in both conditions: 0.06 (n.s.) in the in-sync robot condition and 0.08 (n.s.) in the out-of-sync robot condition.

5.3 Discussion

In Experiment 1, the mixed-rhythm robot condition was marginally different from the unmixed-robot condition; the mixed-rhythm condition exhibited a higher level of synchronization among persons; meaning that H1 was partially supported. Although it was not possible to conclude that the current data successfully verified hypothesis H3, it did not indicate any conflict. This is comparable to results of a study by Nishio [127] that showed a robot hand teleoperated in a synchronized manner to a subject's motion causing the subject to feel like it was an extended body part. Thus, the rubber hand illusion [32] may have occurred between vision and proprioception. In the present chapter, part of the robot body moved in a synchronized manner to a subject's drumbeat in the mixed-rhythm condition. Therefore, it is possible that the subject felt that the robot was an extended body part. When a virtual hand that was recognized as a subject's extended body part is suddenly moved (i.e. asynchronously to the subject), muscle activation of the subject is induced, which matches the asynchronous movement [36]. The robots in the mixed-rhythm condition, which were possibly regarded as extended body parts by the subjects, also exhibited different rhythms from the subjects using other body parts. Therefore, matching the subject's rhythm to that of the subject using a different body part (the experimenter's) was induced. Other experiments suggest that the rubber hand illusion also occurs using a virtual ball, but the illusion is greater when a virtual hand is used [128]. This means that having the perception of body projection can increase the level of synchronization. However, it should be noted that this is not conclusive because the results from the present chapter demonstrated such an increasing trend but reached significance level. Furthermore, it is worth noting that this also means that the rubber hand illusion cannot be considered as the only reason for the increase in the level of synchronization in the experiment. In this experiment, it was also found that the subject felt more enjoyment during the mixed-rhythm condition than during the unmixed one, which supports hypothesis H2. In the unmixed-rhythm condition, the robot lacked synchronous movement with the subject's drumbeats. Thus, when the subject had less chance to observe the synchronicity between the robots and her or himself, the subject felt less enjoyment during this experiment although the influence in the behavioral aspect was only marginally significant. On the other hand, in study 2, when the subject had a chance to observe increased synchrony from more robots, he or she felt more enjoyment although the difference in the behavioral aspect was significant only in the limited pair of conditions. Therefore, in the results from the present chapter, although it cannot be concluded, it seems that the more the subject is synchronized in the group, the more the subject felt enjoyment. One possible reason for such inconclusiveness is that the subjects' flexibility of their drumbeat rhythm or the fragility on the consistency of their drumbeat rhythm were greater than expected. To obtain more conclusive results, a more precise approach to normalize the influence on synchrony by considering the subjects' high flexibility of their drumbeat rhythm or high fragility on the consistency of their drumbeat rhythm is worth investigating. In Experiment 2, it was observed that the tapping interval of the subject in front of an in-sync robot changed compared with when the robot was out-of-sync, which supports hypothesis H3. Meanwhile, the subject felt greater enjoyment in the in-sync condition than in the out-of-sync condition, which supports hypothesis H4 and is consistent with the increase in the level of synchrony. The subjects may have felt that the robot was an extended body part. Slater [36] investigated a virtual hand that a subject recognized as an extended body part, and when it was suddenly moved (i.e., out-of-sync robot), the subject's muscle activation matched asynchronous movement. The in-sync robot condition in this chapter changed the tapping interval of the subject when the robot stopped because of the in-sync movement of the robot. The changes in the subjects' rhythm may be interpreted as resulting from the rubber hand illusion because in both situations the subject appeared to confuse the movements of subject's actual body with those of an in-sync external object. This is comparable to the rubber hand illusion in which synchronous brushing between a subject's actual hand and the rubber hand made the illusion more pronounced. This suggests that the subject's rhythm in Experiment 2 was more synchronized to the experimenter's rhythm because the subject confused the robot's rhythm with their own owing to synchronization to the robot.

5.4 Limitations

In the discussion on Experiment 1, it was argued that the influence of the robots might be enhanced by making the subjects recognize the robots as their extended bodies. It is known that the rubber hand illusion, that is, the illusion that a non-body object is a subject's body, likely occurs when the used non-body obj ect realistically resembles a human hand [129]. This implies that the level of visual similarity among the robots might also influence the current outcome. In this chapter, it was assumed that subjects would have a similar influence as the rubber hand illusion on the robot rhythm when the robot was also in-sync with the subject. It has been argued that the rubber hand illusion [130] and the sense of self-agency [131] are experienced in the synchronization of both the motor signal and visual feedback; this self-agency decreases with the increase in the spatiotemporal discrepancy between the motor signal and visual feedback. Moreover, the self-agency might not be the only factor that influences synchronization. The current chapter has a limitation owing to the experimental design. Since it was used subjects that participated in the study 2 for the first experiment, it should carefully evaluate the result from the viewpoint of subject habituation, fatigue, and the exposure effect. Although it was considered that this was not a serious problem in terms of evaluating the effects of the conditions because it could find significant differences even under such a habituation or positive bias (if there was any) after counter-balancing the order of the conditions, conducting these experiments separately should be considered in the future to increase the reliability of results. In the current experiments, the rhythms of the robots were generated only from the drumbeats of the subject and the experimenter, and the effect on the subject's drumbeats was analyzed. However, it is possible that the subjects' synchronization to the experimenter might be induced also in their different modalities such as neck movements, which the robots used to reproduce the experimenter's rhythm. By considering the increased effect in the mixed rhythm condition where the robots also appeared synchronized to the subject, the effect might be made clearer if the system could represent the subject's potential adaptation of the rhythm in such a different modality.

5.5 Conclusions

In this chapter, it was proposed that multi-party interaction that use group of robots representing two persons' rhythm increased the synchronization in the persons' rhythm. It is possible that the group of robots representing the persons' rhythm induced stronger synchronization than the robots representing only the experimenter's rhythm. However, the synchronized interaction between the robots and people resulted in them enjoying their task more compared with those that solely used the experimenter's rhythm. Additionally, the person was more influence by a robot in-sync with him or her than a robot out-of-sync. The experiments in this chapter focused on robots producing rhythm without using any other social signals, such as emotional expression, eye-contact, or utterances, which are believed to enhance the sense of group belonging and may influence the synchronization it was examined. This influence on the synchronization could be useful in various fields such as musical therapy sessions [132], musical training [133], autism treatment [134], improvement of learning processes [135], and improvement of social skills [136].

CHAPTER 6

STUDY IV: INFLUENCE OF AGENTS IN EMOTIONAL PERCEPTION AND EMOTIONAL CONVEYANCE

In this chapter, it was proposed the use of virtual agents to enhance the degree of perception of a person, who did not clearly was happy, by showing happy virtual agents. Investigations on whether a virtual agent can enhance the degree of the perceived emotion, conveyed emotion, and perceived curvature of the person's lips were performed. In addition, increasing the number of virtual agents and matching their behaviors to enhance emotion perception, emotion conveying, and the perceived curvature of the lips was also proposed in the chapter. Since the virtual agents had similar behavioral cycles when they were matching their behavior, the matching could be considered to be a primitive way of synchronization. Namely, the word synchronization was used for expressing this matching timing behavior.

6.1 Experiment 1

Emotional Avatar Projection System

It was prepare a system using webRTC API for teleconference communication. The system is divided on four processing blocks. It was used a detector for camera frames, a control for the characteristics of the agent, an emotion selector, and a displayed image. The characteristics of the agent to be controlled were emotions, numbers, and position of the agent. The emotions that were available to be displayed by the system were happy, sad, angry, disgust, neutral and fear. The system was able to recognize emotions of the user(experimenter) with a library named face-api.js in real-time. Face-api.js utilize a tensorflow based model developed by Vincent Mühler for the emotion experimenter's recognition. The background on the experiment side was removed by using a machine learning TensorFlow library named bodyPix, posted by Dan Oved, and Tyler Zhu. The emotion recognition accuracy is 95% for happy and between 70% to 80%. The system allowed the user (experimenter) to record different videos for each emotion in order to asynchronously synthesize the virtual agents. The videos were stored on a temporal database. The user was able to select which emotion the virtual agent would display. The user is able to record an exaggerate emotion in order to enhance their own emotion. The video duration was between five to eight seconds. After selecting the emotions, the virtual agent was created and added to the main camera frame. All this processing was doing on a canvas DOM element and controlled by using JavaScript with a 60ms delay for recognizing the emotion. The system block structured is shown on the Figure 6.1.

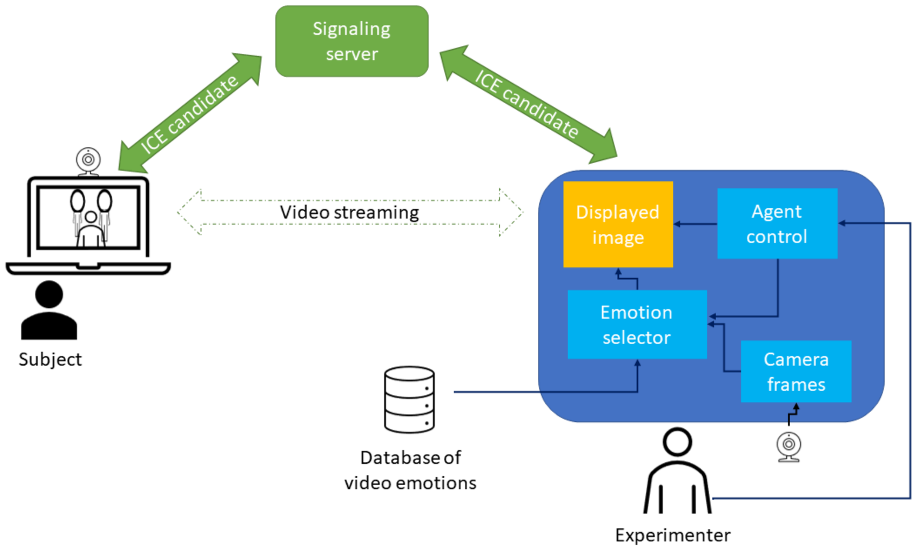

Figure 6.1: System structure for emotion displaying.

In this chapter, it was displayed virtual agents with stronger emotions than the main person. The main focus on the experiment was to explore the basic features of virtual agents in order to convey an emotion that may be difficult to be correctly recognized. Due to this, it was used non real times videos prepared based on our system by using only happy faces. It was used facial features and place them in a thin and small body, making them virtual agents for web conference systems (Figure 6.2).

Figure 6.2: Virtual agent designed for the experiment.

6.1.1 Method

To investigate the effect of a single agent on emotion perception, conveyance of a human, and curvature of the perception of the curvature of lips, an experiment was conducted, wherein two conditions were adopted: no agent condition and one agent condition. Each subject watched the video and were asked to fill on a questionnaire their opinion of the watched video. In this experiment, it was hypothesized that the one agent condition, that is, placing a virtual agent on the shoulder of a person in a video conference environment enhances (H1) the emotion recognition, (H2) emotion conveyance, and (H3) degree of the perceived curvature of the person, better than the no agent condition. The subjects were recruited using a prolific website and provided informed consent at the beginning of the experiment, which was approved by the ethical committee of the Graduate School of Engineering Science, Osaka University.

Subjects

Forty-eight people participated in all the conditions for the first experiment (21 females, 25 males, and 2 people preferred not to disclose their gender, mean age = 25.23, STD=7.38). A few subjects were excluded owing to failure to pass the attention check. The order of the conditions assigned to the subjects was randomly selected such that they were counterbalanced.

Experimental Setup

An experiment was conducted using a website created for this purpose. The videos were downloaded automatically by the website to avoid possible interference owing to poor Internet connections. The answers were stored in a MongoDB database immediately after the participants completed the questionnaire.

Content of Video

One collaborator was recorded with the purpose of saying one stimulus divided on four phrases: “Today is my birthday,” “I am having a birthday party tonight,” “A lot of friends will come,” and “We will have a great time,” The four phrases were said in written order, one after the other. The collaborator recorded these phrases, showing a neutral face and a neutral pitch at the moment of speaking, while looking forward such that the video recorded could be perceived as through the person was in a conference video. The neutral video was recorded in order to verify whether the virtual agent who was recorded with happy emotion reaction to the phrases influence on the emotion perception of neutral collaborator video as happy. The video lasted twelve seconds in total.

Virtual Agent

A virtual agent having the face of a different collaborator and capable of producing emotions by replicating a collaborator's face and body was created. The emotions that the virtual agent produced were represented using only facial and body expressions. The virtual agent did not use sound to express emotion, and its facet was enlarged to facilitate the recognition of the emotion. Further, the virtual agent was placed on the shoulder of the person's video. The virtual agent began by looking towards the front to be perceived as watching the participant who attended the experiment. In the initial state, the virtual agent exhibited a neutral face. Then, he changed his gaze direction to the main person after he/she finished the first phrase. This was followed by looking back at the lens camera after the main person finished the third phrase. The virtual agent began to smile when the person finished saying the second phrase. In Figure 6.3, a sample of the scene of the non-virtual agent condition and one-virtual agent condition are shown.

Figure 6.3: Non-Virtual agent condition and one virtual agent condition

Procedure

The participants were asked to provide their consent to participate in the experiment. They were asked to carefully read the instructions of the experiment and fill out their gender and age information. Following the completion of the video, the participants were able to play the video. They watched the video only once. The video disappeared once watched, and the participants were asked to answer a questionnaire regarding emotion recognition, emotion conveyance, and curvature of the face. Consequently, both conditions were shown, with the questionnaire presented once after each video condition. The conditions were counterbalanced to avoid order effects.

Evaluations

The questionnaire asked six questions to measure emotion recognition, four questions for emotion conveyance, and one question to evaluate the curvature of the face. The questions were rated on a scale with 7 levels, where the middle item was "neither agree neither disagree". Subsequently, indexes found previously in a different experiemt were used to calculate the value for emotion recognition and emotion conveyance, as shown in Table 6.1. It was considered considered invert questions in the questionnaire e.g. "he looked feeling bad" as inverted question of "he looked happy". From the previous studies, the Cronbach's alpha for the variables in Emotion recognition is 0.889 as the Table 6.1 shows. Similarly, the Cronbach's alpha for the variables in Emotion conveyance is 0.840 as the Table 6.1 shows. The indexes for the factors F1 and F2 of the table 6.1 were calculated by using varimax rotation factor analysis type. As this value is high, it was remained the same questionnaire for this experiment.

Table 6.1: Indexes used for each question for measuring emotion recognition and emotion conveyance.

Degree of Emotion recognition α =0.889			
	F1	F2	Communality
He looked fun	.86	.29	.82
He looked energetic	.64	.30	.50
He looked feeling bad*	.47	.28	.30
He looked happy	.88	.16	.80
He looked pleasant	.82	.21	.72
It was difficult to detect which emotion he was expressing*	.62	.25	.44
Degree of conveying emotion α =0.840			
His emotion was easy to understand	.12	.61	.39
His emotion was conveyed to me clearly*	.21	.87	.81
It was difficult for me to understand his feeling*	.40	.63	.56
I think his feeling was explicitly conveyed to me	.35	.72	.64
Eigenvalues	3.55	2.43	5.98
% of variance	35.49	24.33	59.82

Results

The emotion recognition results are shown in Figure 6.4. The average value for the one-agent condition (M=15.98, SD=5.36) was significantly greater than that in the no-agent condition (M=14.87, SD=4.52) ($t(47)=2.08$, $p<.05$, $d=0.22$). The conveyance recognition results are shown in Figure 6.5. In this case, no significant difference was found between the one-agent condition (M=9.53, SD=3.92) and the no-agent condition (M=9.03, SD=3.61) (t (47) =1.31, p=.20, d=0.13). Finally, the curvature results are shown in Figure 6.6 and no significant difference was found between the one-agent condition (M=6.19, SD=0.98) and the no-agent condition (M=6.06, SD=0.60) (t (47) =0.86, p=.39, d=0.16).

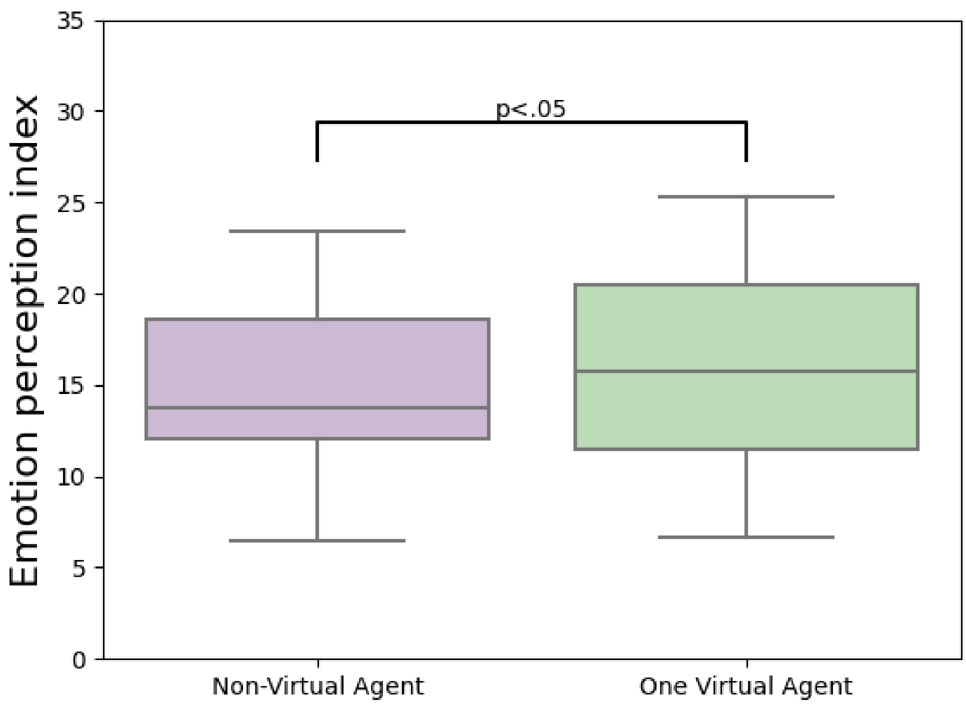

Figure 6.4: Emotion recognition results.

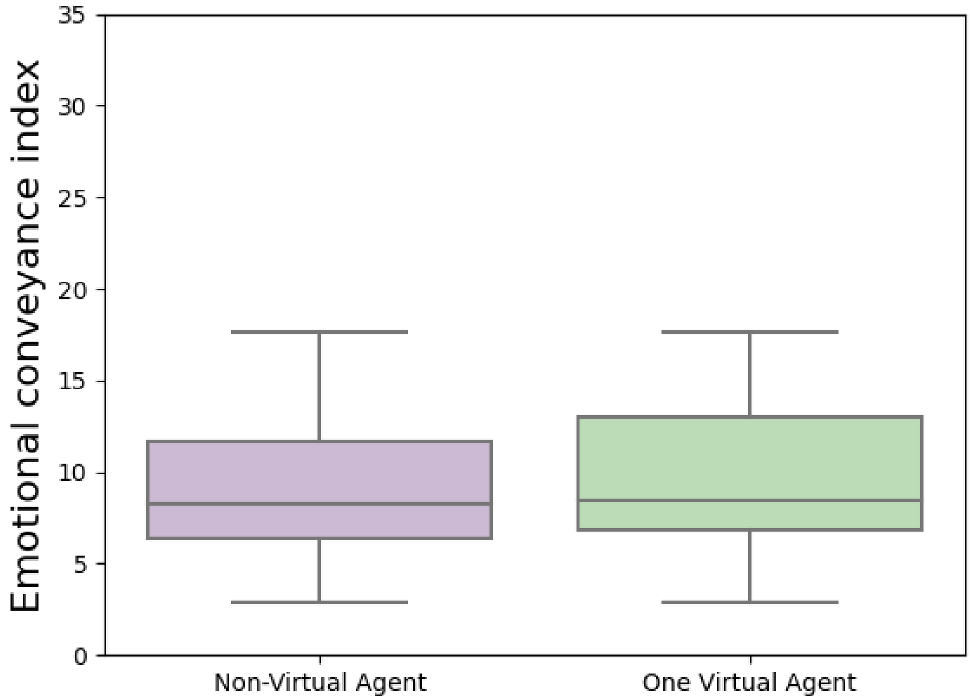

Figure 6.5: Emotion Conveyance results.

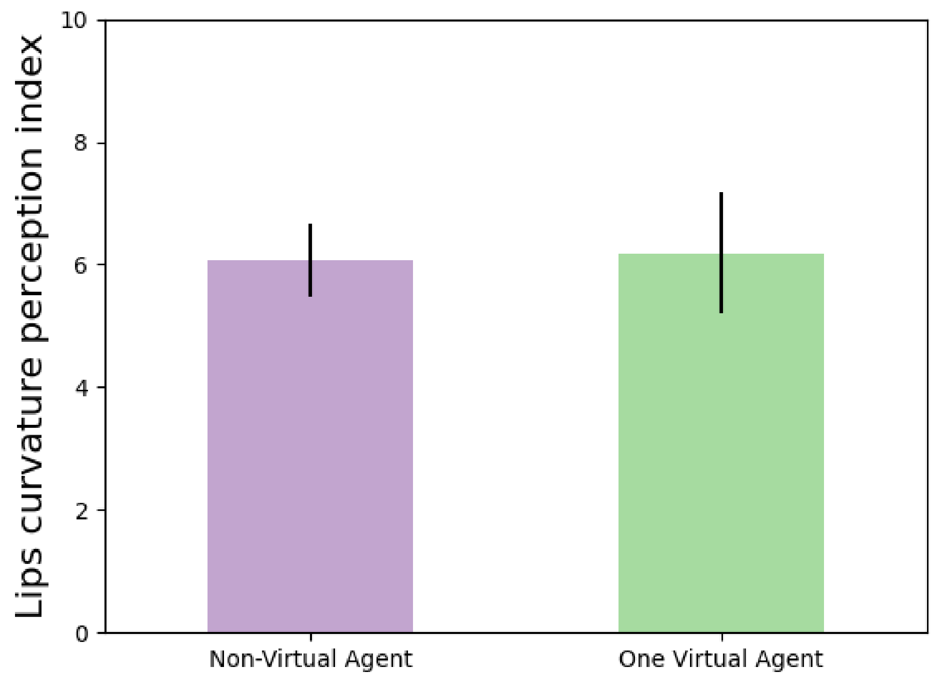

Figure 6.6: Lips Curvature perception result.

6.2 Experiment 2

6.2.1 Method

To investigate the effect of two synchronized agents that can enhance emotion perception, emotion conveyance, and the perception of lip curvature in contrast to a single agent, an experiment was conducted wherein two conditions were adopted: one agent condition and two synchronized agent conditions. In this experiment, it was hypothesized that two synchronized agents, that is, placing two virtual agents on both shoulders of a person in a video conference environment, enhances (H4) the emotion recognition of the person, (H5) the emotion conveyance of the person, and (H6) the degree of the perceived curvature of the person, better than a single agent condition. The subjects were recruited using a prolific website and provided informed consent at the beginning of the experiment, which was approved by the ethical committee of the Graduate School of Engineering Science, Osaka University.

Subjects

Fifty-two people participated in all conditions for the first experiment (22 females and 29 males, mean age = 29.43, STD=8.71). A few subjects were excluded owing to failure to pass the attention check. Moreover, the order of the conditions assigned to the subjects was randomly selected such that they were counterbalanced.

Experimental Setup

An experiment was conducted using a website created for this purpose. The videos were downloaded automatically by the website to avoid possible interference owing to poor Internet connections. Further, the answers were stored in a MongoDB database immediately after the participants completed the questionnaire.

Content of Video

One collaborator was recorded for presenting four phrases: "Today is my birthday," "I am having a birthday party tonight," "A lot of friends will come," and "We will have a great time." The four phrases were said in written order, one after the other. The collaborator recorded these phrases, showing a neutral face and a neutral pitch at the moment of speaking. Moreover, they kept looking forward such that the video recorded could be perceived as though the person was in a conference video. The video lasted 12 s in total.

Virtual Agent

A virtual agent with the face of a different collaborator capable of producing emotions by replicating a collaborator's face and body was created. The emotions that the virtual agents produced were represented using only facial and body expressions. The virtual agent did not use sound to express emotion, and the face of the virtual agent was enlarged to facilitate the recognition of the emotion. Further, they were placed on each shoulder of the person in the video. The virtual agents began by looking towards the front to be perceived as watching the participant who attended the experiment. During the initial state, the virtual agents exhibited a neutral face. Then they changed their gaze direction to the main person after he/she finished saying the first phrase and looked back at the lens camera after the main person finished the third phrase. The virtual agents started to smile when the person finished saying the second phrase: The second virtual agent was a copy of the first virtual agent, synchronizing its behavior with the timing. In Figure 6.7, a sample of the scene of the two virtual agent conditions and the one-virtual agent condition are shown.

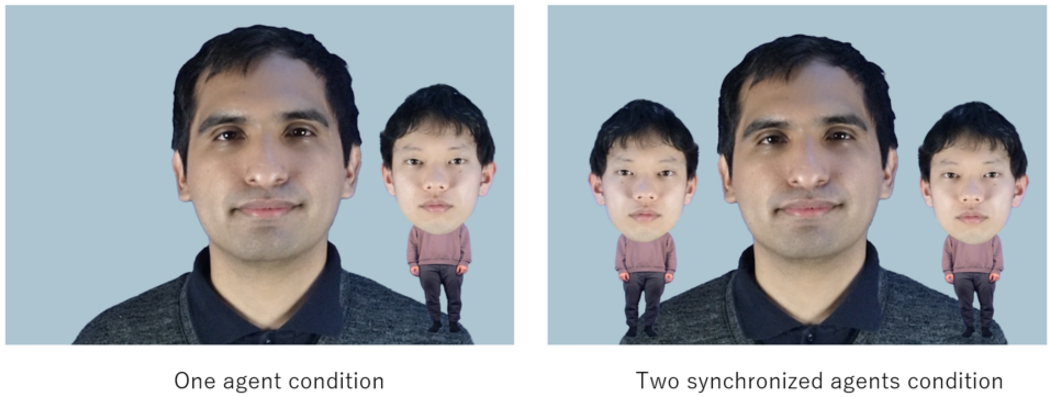

Figure 6.7: Lips Curvature perception result.

Procedure

The participants were asked to provide their consent to participate in the experiment. They were asked to carefully read the instructions of the experiment and fill out their gender and age information. After the video was completed, the participants were able to play the video. Participants watched the video only once. The video disappeared once it was watched, and the participants were asked to answer a questionnaire regarding emotion recognition, emotion conveyance, and curvature of the face.

Evaluations

The questionnaire asked six questions to measure emotion recognition, four questions for emotion conveyance, and one question to evaluate the curvature of the face. The questions were rated on a scale with 7 levels, where the neutral perception was considered in the middle item of all the questions that was "neither agree neither disagree".

Results

The emotion recognition results are shown in Figure 6.8. The average value for the two synchronized agents' condition (M=17.60, SD=4.83) was significantly greater than that of the no-agent condition (M=16.69, SD=4.91) ($t(51) = 2.14$, $p < .05$, $d = 0.19$). The conveyance recognition results are shown in Figure 6.9. In this case, the average value for the two synchronized agents' condition (M=10.76, SD=4.15) was marginally greater than that for the one-agent condition (M=10.06, SD=3.74) ($t(51) = 1.86$, $p = .07$, $d = 0.18$). Finally, the curvature results are shown in Figure 6.10, and no significant difference between the two synchronized agents' conditions (M=6.29, SD=0.90) and the one-agent condition (M=6.20, SD=0.78) ($t(51) = 1.00$, $p = .32$, $d = 0.11$) was observed.

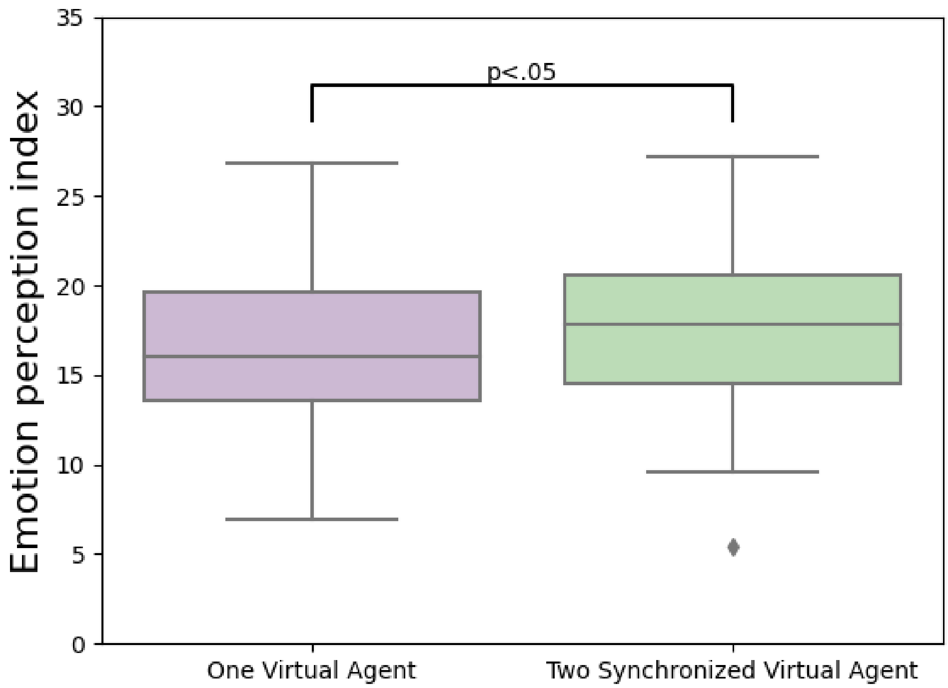

Figure 6.8: Emotion recognition results.

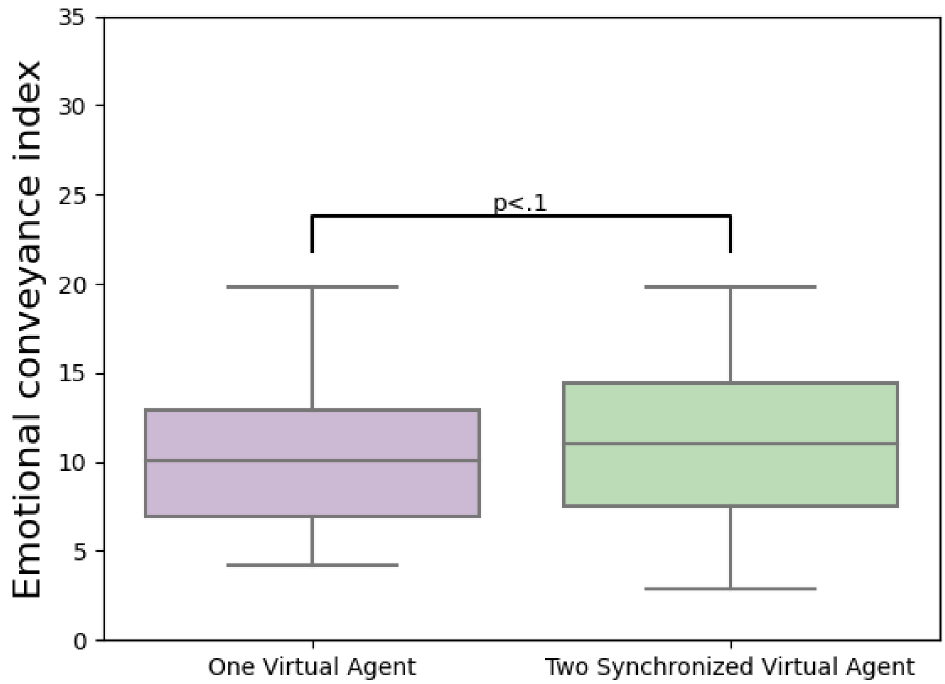

Figure 6.9: Emotion Conveyance results.

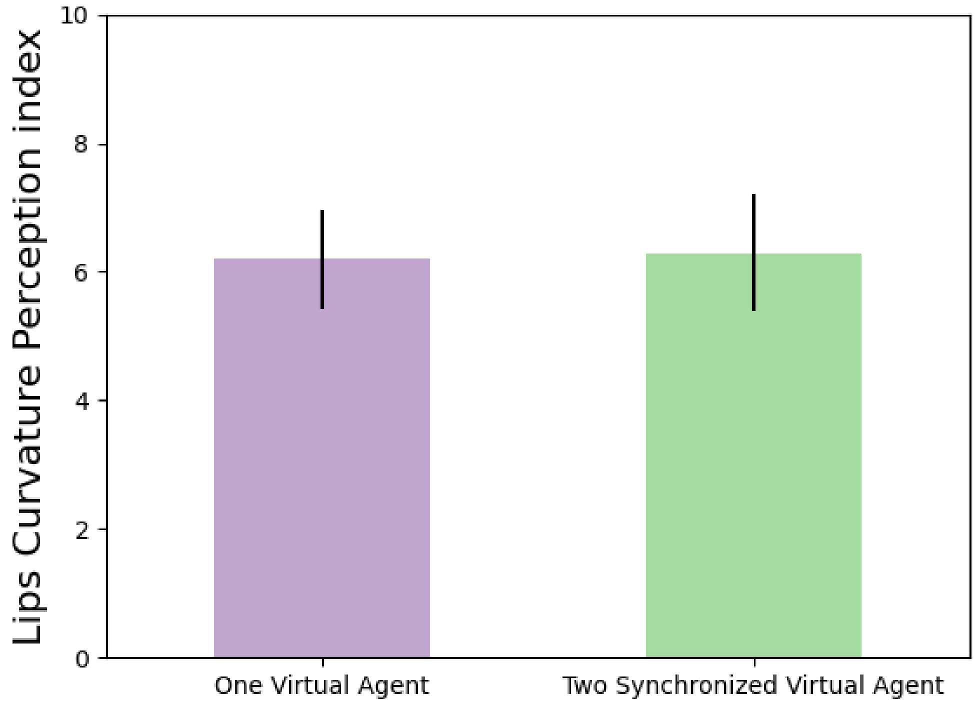

Figure 6.10: Lips Curvature perception result.

6.3 Discussion

The level of perceived emotion reported in Experiment 1 suggests that the one virtual agent condition enhances the perceived emotion of a person more than the non-virtual agent condition. However, no significant difference was observed for emotion conveyance or the curvature of the lips. In other words, H1 was supported but H2 and H3 were not. For H1, humans recognize easily happy facial expressions in groups of people [137]. Moreover, a body with a posture that shows emotions similar to that of facial expressions aids in recognizing emotions [138]. Thus, showing a happy representation of the face and body of a person near a main person without a clear expression can enhance emotion recognition of this emotion. In these experiment, it was added a virtual agent with the face and body characteristics of the emotion that needed to be conveyed to enhance the emotion recognition (H1) of a person. In case of H2, it is known that agents in groups can convey emotions more easily than isolated agents [139]. The lack of a significant difference in the conveyance of the emotion(H2) may be attributed to the fact that the main person and the virtual agent were not considered as a single group, but as isolated agents. Therefore, the influence of the virtual agent on enhancing the level of emotional conveyance of the main person may be weak and insufficient. Whereas, regarding H3, multiple faces with a laughter background were used to increase the perception of lip curvature [97]. In the experiment conducted, only a single agent was implemented; therefore, the effect of the number of virtual agents must be studied in detail in future experiments.

The level of perceived emotion reported in Experiment 2 (H4) suggests that two synchronized virtual agents enhance the perceived emotion of a person better than a single agent. In addition, a marginally significant difference between the two synchronized virtual agents and one virtual agent condition (H5) was observed as well. However, no significant difference in the lip curvature (H5) was obtained. In other words, H4 was supported, but H5 and H6 were not. For H4, it is known that increasing the number of synchronized agents and rendering them as part of a group enhances enjoyment perception [52]. It is believed that the group of two agents enhanced the emotion recognition of the main person better than the single agent condition (H4) because the virtual agents were perceived as a group because of their synchronization. IN case of H5, finding a marginally significant difference in conveying emotions may support the idea that increasing the number of agents on a group of more than two virtual agents could enhance the emotion conveyance of the main person compared to a single virtual agent. However, for H6, it is necessary to study the influence of the group of more synchronized agents in more detail in future experiments.

6.4 Limitations

Although this experiment was not interactive, the participants were asked to watch the videos offline making possible that people's perceptions differ in real-time interactions. Moreover, as the experiment was not conducted on-line (real-time), it was not possible to measure the probable delays, the manual control of the agent and the correct method to use these agents on online conference systems. Moreover, the synchronization described on this experiment consisted on matching the behavior of the virtual agents. The current features studied about synchronization might be insuffient and new features might be added e.g., emotion timing, or voice onset synchronization, etc. In the discussion of Experiment 1, the effect of one agent was confirmed; however, the effective elements need to be studied more precisely, such as the effect of body, size of agent, size of face (in the current experiment, the face was not proportional to the body), and the position of the agent. In Experiment 2, two time-synchronized virtual agents were evaluated. In future studies, the effective degree of synchronization must be studied. Moreover, because autonomy and spontaneous movements of agents have been reported as factors that increase the agency and consequently the positive feeling of the participant regarding it [140], completely synchronized behavior may not be sufficient for long-term interaction and real-world application. The effect of the number of agents is an additional factor that needs to be studied. As the number of agents during the communication has a remarkable effect on smoothening the interaction, improving the feeling of autonomy and perception regarding involvement in the conversation should have a positive effect on enhancing recognition; such an effect in the case of our virtual agent still needs to be explored. In addition, the effect of other synchronous types of agents, for example, the synchrony of the geometrical component features of the agents, the synchronization in the emotional state, and the synchronization of the attention behavior, must be researched in future works. Nevertheless, this chapter conducted experiments which were focused only on the emotional expressions of happiness. Moreover, it was intentionally not included a scale for measuring neutral emotion as the main intention was to seem the main person on the video look happier. Thus, precise, and systematic research on other emotions, such as those presented and classified in the circumplex model of emotion by Russell [141] or basic universal emotions by Ekman [142] is required in future studies.

6.5 Conclusion

In this chapter, it was proposed that the presence of virtual agents in videos can enhance the perceived emotion of a main person even when the person has a neutral face. A virtual agent containing an exaggerated collaborator's face and a shrunken body was developed, which was located on the shoulder of the main person in a video-conferencing environment. In addition, videos wherein the agents were presented and behaved synchronously were prepared. To verify the proposed method, two experiments utilizing human subjects were conducted to explore two comparisons: 1) the effect of utilizing the virtual agent compared to no virtual agent, and 2) the effect of utilizing two synchronized virtual agents compared to one virtual agent. In these experiments, improvement in the perceived emotion expression, emotion conveyance, and the degree of the smile of the main person in terms of perceived level of curve in the mouth of the main person were evaluated, as the experiments was conducted for the case of the emotion of happiness. The results of the statistical hypothesis test indicated that one agent improved emotion recognition and conveyance compared to the case in which there was no agent. In addition, it was revealed that two synchronized agents performed better than one agent in terms of emotion recognition. These results show the potential of the proposed virtual agent in improving the online communication of humans by increasing their positive perceptions of each other. However, these experiments did not verify the effect of different shapes of the agent or different types of synchronization, which will be analyzed in future experiments. Moreover, the current evaluation was performed by only checking the effect of watching prerecorded videos, and future studies will use the designed system for real-time communication experiments. Moreover, this chapter focused on virtual agents that produce only happy behavior without using pitch voice, topic context regarding the smile, or having a background which are believed to enhance the emotion recognition more.

CHAPTER 7

CONCLUSIONS

The purpose of the studies mentioned was to find the influential characteristics that agents have in humans. For this, it was decided to explore the influence of agents in synchronization, movement interference, and emotional perception. it was tested the influence of small humanoid robots, projected big size robots, and virtual humanoid agents throughout the experiment. The information representation of humans on robots and its influence on their behavior was also tested. Moreover, whether a human-like behavior is more influential than robot-like behavior was evaluated. In addition, whether an agent is able to influence others' emotional perception and conveyance of a person was evaluated. In other words, it was explore the influence of a group of agents as a catalyst for behavioral and emotional influence on humans.

To reach this goal, it was explore three studies in which the influence of three different agent embodiment was used. Therefore, in this dissertation, a full chapter was assigned to present each one of these influences separately, and below is the summary:

CHAPTER 3 presented a group of projected agents for interference with the hand movement of people and for synchronize the agents in antiphase synchronization. The projected agents have two different avatars: a) the avatar was similar to the first experiment avatar, and b) the avatar was a human recorded previously. In the experiment, the avatars move at two different speeds: a)linear speed which was matched to the speed of the robot, and b) biological speed which was matched to the speed of the human. The robot agent was programmed to move at the same speed as the human, and the human avatar -

which was recorded from a video was processed to have the same speed as the robot. The projected agents were increased for testing the group influence in the humans.

From the results of the experiments, three agents were found to be more influential in the hand movement interference than a single agent. Moreover, using human avatars was found to be more influential in the antiphase synchronization of the human with the agents than using robot avatars.

CHAPTER 4 presented a group of humanoid small robots -normally used for enhancing communication- for enhancing synchronization between two persons. The robots have big eyes and they were supposed to move accordingly to musical patterns generated by a human or by a computer. The goal of the chapter was to explore the influence of the robots on humans synchronization. In order to explore this influence, three different experiments were conducted for evaluating the number of robots, the rhythm patterns shared on the robots, and the influence of a single robot on humans' rhythmic patterns. Moreover, the enjoyment that humans experience after being involved in the interaction with the robots was explored as well.

From the results of the experiments, three robots were found to be more influential in synchronizing two humans than having the humans alone playing the drums. The results suggested that a group of robots are more influential on synchronization than having no robots.

CHAPTER 5 presented a robot and a group of robots that shared the rhythm of two people into the robots. It was found that sharing the robots' bodies to show people rhythm was more influential than not sharing the robots' bodies to show people rhythm. Moreover, having the robot in synchronization with a person influenced the person's perception of the robot.

From the result, it is suggested that sharing robots' bodies to show people's rhythm influences the perception of the robot and the synchronization of the person to the robot.

CHAPTER 6 demonstrates the possible effect of using virtual agents on real-life applications. The agents used in this experiment were prepared for enhancing emotion perception and emotion conveyance. In the experiments, the virtual agents were tested by increasing the number and synchronizing them.

From the results of the experiments, having one virtual agent enhances the emotional perception more than having no virtual agents. Moreover, having two synchronous virtual agents enhances the emotional perception more than having one virtual agent. The emotional conveyance might be influenced by two synchronous virtual agents more than having one single virtual agent, but more experiments need to be conducted in order to verify this fact.

In general, the influence of a group of agents was found in simple synchronous simple tasks, such as playing drums. Moreover, the group of agents was found to be more influential in human behavior interference, such as moving the hand. In more goal-oriented tasks, such as synchronizing on purpose, the influential of the avatar is more relevant than the group. Because of this, the influence of the avatar of agents in the emotional perception is more influential as the task is complex, but the influence on the conveyance of the number of agents might be more influential.

ACKNOWLEDGMENTS

First and foremost, I would like to express my honest gratitude to Professor Hiroshi Ishiguro for his advice and guidance throughout my studies. I would also like to thank my supervisors Professor Yuichiro Yoshikawa and Dr. Hamed Mahzoon, for their patience, dedication, and comprehension in every step of this project. Their huge support helped me to overcome the obstacles I faced throughout my studies. Similarly, I would like to express my gratitude to my wife who supports me emotionally on every step of this journey and my daughter who was my biggest motivation to concrete this project. Also, I would like to express my gratitude to my parents and sister for their encouraging words. I wanted to express my gratitude to Dr. Eduardo Castello and Dr. Fabio dalla libera for their positive feedback on my work and their wise advice.

BIBLIOGRAPHY

[1] Arkady Pikovsky, Michael Rosenblum, and Jürgen Kurths. Synchronization: a universal concept in nonlinear science, 2002.

[2] JA Scott Kelso and AJ van Opstal. Dynamic patterns: The self-organization of brain and behavior. *Journal of Cognitive Neuroscience*, 8(4):385, 1996.

[3] Masahiro Kawasaki, Yohei Yamada, Yosuke Ushiku, Eri Miyauchi, and Yoko Yamaguchi. Inter-brain synchronization during coordination of speech rhythm in human-to-human social interaction. *Scientific reports*, 3(1):1–8, 2013.

[4] Michael J Richardson, Kerry L Marsh, Robert W Isenhower, Justin RL Goodman, and Richard C Schmidt. Rocking together: Dynamics of intentional and unintentional interpersonal coordination. *Human movement science*, 26(6):867–891, 2007.

[5] Niek R van Ulzen, Claudine JC Lamoth, Andreas Daffertshofer, Gün R Semin, and Peter J Beek. Characteristics of instructed and uninstructed interpersonal coordination while walking side-by-side. *Neuroscience letters*, 432(2):88–93, 2008.

[6] Michael J Richardson, Kerry L Marsh, and RC Schmidt. Effects of visual and verbal interaction on unintentional interpersonal coordination. *Journal of Experimental Psychology: Human Perception and Performance*, 31(1):62, 2005.

[7] Richard C Schmidt, Claudia Carello, and Michael T Turvey. Phase transitions and critical fluctuations in the visual coordination of rhythmic movements between people. *Journal of experimental psychology: human perception and performance*, 16(2):227, 1990.

[8] Joseph Jaffe, Beatrice Beebe, Stanley Feldstein, Cynthia L Crown, Michael D Jasnow, Philippe Rochat, and Daniel N Stern. Rhythms of dialogue in infancy: Coordinated timing in development. *Monographs of the society for research in child development*, pages i–149, 2001.

[9] Piercarlo Valdesolo, Jennifer Ouyang, and David DeSteno. The rhythm of joint action: Synchrony promotes cooperative ability. *Journal of experimental social psychology*, 46(4):693–695, 2010.

[10] Lynden K Miles, Louise K Nind, and C Neil Macrae. The rhythm of rapport: Interpersonal synchrony and social perception. *Journal of experimental social psychology*, 45(3):585–589, 2009.

[11] Alexander Mörtl, Tamara Lorenz, and Sandra Hirche. Rhythm patterns interaction-synchronization behavior for human-robot joint action. *PloS one*, 9(4):e95195, 2014.

[12] Paula Marie Theresa Scothern. *The music-archaeology of the Palaeolithic within its cultural setting.* PhD thesis, University of Cambridge, 1993.

[13] Björn Merker. Synchronous chorusing and the origins of music. *Musicae Scientiae*, 3(1_suppl):59–73, 1999.

[14] Steven Brown. Evolutionary models of music: From sexual selection to group selection. In *Perspectives in ethology*, pages 231–281. Springer, 2000.

[15] William Weber. Keeping together in time: Dance and drill in human history. *The Journal of Interdisciplinary History*, 27(4):664–666, 1997.

[16] Juan G Roederer. The search for a survival value of music. *Music perception*, 1(3):350–356, 1984.

[17] Zoltan Neda, Erzsébet Ravasz, Yves Brechet, Tamás Vicsek, and A-L Barabási. The sound of many hands clapping: Tumultuous applause can transform itself into waves of synchronized clapping. *Nature (London)*, 403(6772):849–850, 2000.

[18] Tuomas Eerola, Geoff Luck, Petri Toiviainen, et al. An investigation of pre-schoolers' corporeal synchronization with music. In *Proceedings of the 9th international conference on music perception and cognition*, pages 472–476. Citeseer, 2006.

[19] Marek P Michalowski, Selma Sabanovic, and Hideki Kozima. A dancing robot for rhythmic social interaction. In *Proceedings of the ACM/IEEE international conference on Human-robot interaction*, pages 89–96, 2007.

[20] Marek P Michalowski, Reid Simmons, and Hideki Kozima. Rhythmic attention in child-robot dance play. In *RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication*, pages 816–821. IEEE, 2009.

[21] Eleanor Avrunin, Justin Hart, Ashley Douglas, and Brian Scassellati. Effects related to synchrony and repertoire in perceptions of robot dance. In *Proceedings of the 6th international conference on Human-robot interaction*, pages 93–100, 2011.

[22] Guy Hoffman and Keinan Vanunu. Effects of robotic companionship on music enjoyment and agent perception. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pages 317–324. IEEE, 2013.

[23] Christopher Crick, Matthew Munz, and Brian Scassellati. Synchronization in social tasks: Robotic drumming. In *ROMAN 2006-The 15th IEEE international symposium on robot and human interactive communication*, pages 97–102. IEEE, 2006.

[24] Hatice Kose-Bagci, Kerstin Dautenhahn, and Chrystopher L Nehaniv. Emergent dynamics of turn-taking interaction in drumming games with a humanoid robot. In *RO-MAN 2008-The 17th IEEE International Symposium on Robot and Human Interactive Communication*, pages 346–353. IEEE, 2008.

[25] Tariq Iqbal, Maryam Moosaei, and Laurel D Riek. Tempo adaptation and anticipation methods for human-robot teams. In *RSS, Planning HRI: Shared Autonomy Collab. Robot. Workshop*, 2016.

[26] Jiří Mates. A model of synchronization of motor acts to a stimulus sequence. *Biological cybernetics*, 70(5):463–473, 1994.

[27] Kyohei Tatsukawa, Tamami Nakano, Hiroshi Ishiguro, and Yuichiro Yoshikawa. Eyeblink synchrony in multimodal human-android interaction. *Scientific reports*, 6(1):1–8, 2016.

[28] James M Kilner, Yves Paulignan, and Sarah-Jayne Blakemore. An interference effect of observed biological movement on action. *Current biology*, 13(6):522–525, 2003.

[29] Thierry Chaminade, David W Franklin, Erhan Oztop, and Gordon Cheng. Motor interference between humans and humanoid robots: Effect of biological and artificial motion. In *Proceedings. The 4th International Conference on Development and Learning, 2005*, pages 96–101. IEEE, 2005.

[30] Christopher Crick, Matthew Munz, Tomislav Nad, and Brian Scassellati. Robotic drumming: Synchronization in social tasks. In *IEEE ROMAN*, pages 97–102, 2006.

[31] Tomio Watanabe, Masashi Okubo, and Hiroki Ogawa. An embodied interaction robots system based on speech. In *8th IEEE International Workshop on Robot and Human Interaction. RO-MAN'99 (Cat. No. 99TH8483)*, pages 225–230. IEEE, 1999.

[32] Matthew Botvinick and Jonathan Cohen. Rubber hands 'feel' touch that eyes see. *Nature*, 391(6669):756–756, 1998.

[33] Manos Tsakiris and Patrick Haggard. The rubber hand illusion revisited: visuotactile integration and self-attribution. *Journal of experimental psychology: Human perception and performance*, 31(1):80, 2005.

[34] Maryam Alimardani, Shuichi Nishio, and Hiroshi Ishiguro. Humanlike robot hands controlled by brain activity arouse illusion of ownership in operators. *Scientific reports*, 3(1):1–5, 2013.

[35] Maryam Alimardani, Shuichi Nishio, and Hiroshi Ishiguro. Effect of biased feedback on motor imagery learning in bci-teleoperation system. *Frontiers in systems neuroscience*, 8:52, 2014.

[36] Mel Slater, Daniel Pérez Marcos, Henrik Ehrsson, and Maria V Sanchez-Vives. Towards a digital body: the virtual arm illusion. *Frontiers in human neuroscience*, page 6, 2008.

[37] S Milgram. Behavioral study of obedience, j. abnormz. soc, 1963.

[38] Michael A Wallach, Nathan Kogan, and Daryl J Bem. Group influence on individual risk taking. *The Journal of Abnormal and Social Psychology*, 65(2):75, 1962.

[39] Geoffrey L Cohen. Party over policy: The dominating impact of group influence on political beliefs. *Journal of personality and social psychology*, 85(5):808, 2003.

[40] Solomon E Asch. Opinions and social pressure. *Scientific American*, 193(5):31–35, 1955.

[41] David A Wilder. Perception of groups, size of opposition, and social influence. *Journal of Experimental Social Psychology*, 13(3):253–268, 1977.

[42] Pei-Luen Patrick Rau, Ye Li, and Jun Liu. Effects of a social robot's au-

tonomy and group orientation on human decision-making. *Advances in Human-Computer Interaction*, 2013, 2013.

[43] Gordon Briggs and Matthias Scheutz. How robots can affect human behavior: Investigating the effects of robotic displays of protest and distress. *International Journal of Social Robotics*, 6(3):343–355, 2014.

[44] Tsunehiro Arimoto, Yuichiro Yoshikawa, and Hiroshi Ishiguro. Multiple-robot conversational patterns for concealing incoherent responses. *International Journal of Social Robotics*, 10(5):583–593, 2018.

[45] Takamasa Iio, Yuichiro Yoshikawa, and Hiroshi Ishiguro. Pre-scheduled turn-taking between robots to make conversation coherent. In *Proceedings of the Fourth International Conference on Human Agent Interaction*, pages 19–25, 2016.

[46] Stanley Milgram and Christian Gudehus. Obedience to authority, 1978.

[47] Masahiro Shiomi and Norihiro Hagita. Do synchronized multiple robots exert peer pressure? In *Proceedings of the fourth international conference on human agent interaction*, pages 27–33, 2016.

[48] Takayuki Kanda, Rumi Sato, Naoki Saiwaki, and Hiroshi Ishiguro. A two-month field trial in an elementary school for long-term human–robot interaction. *IEEE Transactions on robotics*, 23(5):962–971, 2007.

[49] Yuicho Yoshikawa, Takamasa Iio, Tsunehiro Arimoto, Hiroaki Sugiyama, and Hiroshi Ishiguro. Proactive conversation between multiple robots to improve the sense of human–robot conversation. In *2017 AAAI Fall Symposium Series*, 2017.

[50] Eri Takano, Yoshio Matsumoto, Yutaka Nakamura, Hiroshi Ishiguro, and Kazuomi Sugamoto. Psychological effects of an android bystander on human-human communication. In *Humanoids 2008-8th IEEE-RAS International Conference on Humanoid Robots*, pages 635–639. IEEE, 2008.

[51] Michihiro Shimada, Yuichiro Yoshikawa, and Hiroshi Ishiguro. Social balancing effect of eye contact. In *RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication*, pages 1119–1124. IEEE, 2009.

[52] Alexis Meneses, Yuichiro Yoshikawa, and Hiroshi Ishiguro. Effect of syn-

chronous robot motion on human synchrony and enjoyment perception. *Interaction Studies*, 22(1):86–109, 2021.

[53] Minoru Asada. Towards artificial empathy. *International Journal of Social Robotics*, 7(1):19–33, 2015.

[54] Takahiro Yokozuka, Eisuke Ono, Yuki Inoue, Ken-Ichiro Ogawa, and Yoshihiro Miyake. The relationship between head motion synchronization and empathy in unidirectional face-to-face communication. *Frontiers in psychology*, 9:1622, 2018.

[55] Hiroshi Ishiguro. Studies on humanlike robots–humanoid, android and geminoid. In *International Conference on Simulation, Modeling, and Programming for Autonomous Robots*, pages 2–2. Springer, 2008.

[56] Christian Becker-Asano, Kohei Ogawa, Shuichi Nishio, and Hiroshi Ishiguro. Exploring the uncanny valley with geminoid hi-1 in a real-world application. In *Proceedings of IADIS International conference interfaces and human computer interaction*, pages 121–128, 2010.

[57] Daisuke Sakamoto and Hiroshi Ishiguro. Geminoid: Remote-controlled android system for studying human presence. *Kansei Engineering International*, 8(1):3–9, 2009.

[58] Mel Slater, D-P Pertaub, and Anthony Steed. Public speaking in virtual reality: Facing an audience of avatars. *IEEE Computer Graphics and Applications*, 19(2):6–9, 1999.

[59] Guo Freeman, Samaneh Zamanifard, Divine Maloney, and Alexandra Adkins. My body, my avatar: How people perceive their avatars in social virtual reality. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, pages 1–8, 2020.

[60] Harrison Jesse Smith and Michael Neff. Communication behavior in embodied virtual reality. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, pages 1–12, 2018.

[61] Mohammad Bani Younes and Samer Al-Zoubi. The impact of technologies on society: A review. *IOSR Journal of Humanities and Social Science*, 20(2):82–86, 2015.

[62] Sujatha Krishnan-Barman, Paul AG Forbes, and Antonia F de C Hamil-

ton. How can the study of action kinematics inform our understanding of human social interaction? *Neuropsychologia*, 105:101–110, 2017.

[63] Jennifer Cook, David Swapp, Xueni Pan, Nadia Bianchi-Berthouze, and Sarah-Jayne Blakemore. Atypical interference effect of action observation in autism spectrum conditions. *Psychological medicine*, 44(4):731–740, 2014.

[64] Anna Samira Praetorius and Daniel Görlich. How avatars influence user behavior: a review on the proteus effect in virtual environments and video games. In *International Conference on the Foundations of Digital Games*, pages 1–9, 2020.

[65] Ivelina V Piryankova, Jeanine K Stefanucci, Javier Romero, Stephan De La Rosa, Michael J Black, and Betty J Mohler. Can i recognize my body's weight? the influence of shape and texture on the perception of self. *ACM Transactions on Applied Perception (TAP)*, 11(3):1–18, 2014.

[66] Erhan Oztop, David W Franklin, Thierry Chaminade, and Gordon Cheng. Human–humanoid interaction: Is a humanoid robot perceived as a human? *International Journal of Humanoid Robotics*, 2(04):537–559, 2005.

[67] Michael E Walker, Daniel Szafir, and Irene Rae. The influence of size in augmented reality telepresence avatars. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 538–546. IEEE, 2019.

[68] David Weibel, Daniel Stricker, Bartholomäus Wissmath, and Fred W Mast. How socially relevant visual characteristics of avatars influence impression formation. *Journal of Media Psychology: Theories, Methods, and Applications*, 22(1):37, 2010.

[69] Li-Ann Leow, Kristina Waclawik, and Jessica A Grahn. The role of attention and intention in synchronization to music: Effects on gait. *Experimental Brain Research*, 236(1):99–115, 2018.

[70] Shinya Fujii and Gottfried Schlaug. The harvard beat assessment test (h-bat): a battery for assessing beat perception and production and their dissociation. *Frontiers in human neuroscience*, 7:771, 2013.

[71] Maurice Mohr, Marius Nann, Vinzenz von Tscharner, Bjoern Eskofier, and Benno Maurus Nigg. Task-dependent intermuscular motor unit synchronization between medial and lateral vastii muscles during dynamic and isometric squats. *PloS one*, 10(11):e0142048, 2015.

[72] Ailin Leng, Lana Friesen, Kenan Kalayci, and Priscilla Man. A minimum effort coordination game experiment in continuous time. *Experimental Economics*, 21(3):549–572, 2018.

[73] Anton Stepanov, Andrey Lange, Nikita Khromov, Alexander Korotin, Evgeny Burnaev, and Andrey Somov. Sensors and game synchronization for data analysis in esports. In *2019 IEEE 17th International Conference on Industrial Informatics (INDIN)*, volume 1, pages 933–938. IEEE, 2019.

[74] Kiavash Bahreini, Rob Nadolski, and Wim Westera. Communication skills training exploiting multimodal emotion recognition. *Interactive Learning Environments*, 25(8):1065–1082, 2017.

[75] Tassilo Momm, Gerhard Blickle, Yongmei Liu, Andreas Wihler, Mareike Kholin, and Jochen I Menges. It pays to have an eye for emotions: Emotion recognition ability indirectly predicts annual income. *Journal of Organizational Behavior*, 36(1):147–163, 2015.

[76] Kurt Hugenberg and John Paul Wilson. Faces are central to social cognition. 2013.

[77] Paula M Niedenthal and Markus Brauer. Social functionality of human emotion. *Annual review of psychology*, 63(1):259–285, 2012.

[78] Paul Ekman. Are there basic emotions? 1992.

[79] Ralph Adolphs. Recognizing emotion from facial expressions: psychological and neurological mechanisms. *Behavioral and cognitive neuroscience reviews*, 1(1):21–62, 2002.

[80] Ralph Adolphs, Frederic Gosselin, Tony W Buchanan, Daniel Tranel, Philippe Schyns, and Antonio R Damasio. A mechanism for impaired fear recognition after amygdala damage. *Nature*, 433(7021):68–72, 2005.

[81] Baron-Cohen, Simon and Wheelwright, Sally and Jolliffe, and Therese. Is there a" language of the eyes"? evidence from normal adults, and adults with autism or asperger syndrome. *Visual cognition*, 4(3):311–331, 1997.

[82] Shelly L Gable, Gian C Gonzaga, and Amy Strachman. Will you be there for me when things go right? supportive responses to positive event disclosures. *Journal of personality and social psychology*, 91(5):904, 2006.

[83] Juan Song, Yanqiu Wei, and Han Ke. The effect of emotional information from eyes on empathy for pain: A subliminal erp study. *PLoS One*, 14(12):e0226211, 2019.

[84] Robert H Frank. *Passions within reason: The strategic role of the emotions.* WW Norton & Co, 1988.

[85] Iain R Murray and John L Arnott. Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *The Journal of the Acoustical Society of America*, 93(2):1097–1108, 1993.

[86] Paul Ed Ekman and Richard J Davidson. *The nature of emotion: Fundamental questions.* Oxford University Press, 1994.

[87] Charles Darwin and Phillip Prodger. *The expression of the emotions in man and animals*. Oxford University Press, USA, 1998.

[88] Eva G Krumhuber, Sylwia Hyniewska, and Anna Orlowska. Contextual effects on smile perception and recognition memory. *Current Psychology*, pages 1–9, 2021.

[89] Justin P Friesen, Kerry Kawakami, Larissa Vingilis-Jaremko, Regis Caprara, David M Sidhu, Amanda Williams, Kurt Hugenberg, Rosa Rodríguez-Bailón, Elena Cañadas, and Paula Niedenthal. Perceiving happiness in an intergroup context: The role of race and attention to the eyes in differentiating between true and false smiles. *Journal of Personality and Social Psychology*, 116(3):375, 2019.

[90] Peter J Reschke and Eric A Walle. The unique and interactive effects of faces, postures, and scenes on emotion categorization. *Affective Science*, 2(4):468–483, 2021.

[91] Sally Ahmed Mosad Eltenahy. Facial recognition and emotional expressions over video conferencing based on web real time communication and artificial intelligence. In *Enabling Machine Learning Applications in Data Science*, pages 29–37. Springer, 2021.

[92] CNW Geraets, S Klein Tuente, BP Lestestuiver, M Van Beilen, SA Nijman, JBC Marsman, and W Veling. Virtual reality facial emotion recognition in social environments: An eye-tracking study. *Internet Interventions*, 25:100432, 2021.

[93] Hanneke KM Meeren, Corné CRJ van Heijnsbergen, and Beatrice de Gelder. Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences*, 102(45):16518–16523, 2005.

[94] Daichi Shimizu and Takeshi Okada. Synchronization and coordination of art performances in highly competitive contexts: Battle scenes of expert breakdancers. *Frontiers in Psychology*, 12:635534, 2021.

[95] Kai Spiegelhalder, Sabine Ohlendorf, Wolfram Regen, Bernd Feige, Ludger Tebartz van Elst, Cornelius Weiller, Jürgen Hennig, Mathias Berger, and Oliver Tüscher. Interindividual synchronization of brain activity during live verbal communication. *Behavioural brain research*, 258:75–79, 2014.

[96] Daniël Lakens and Mariëlle Stel. If they move in sync, they must feel in sync: Movement synchrony leads to attributions of rapport and entitativity. *Social Cognition*, 29(1):1, 2011.

[97] Aleksandra Sherman, Timothy D Sweeny, Marcia Grabowecky, and Satoru Suzuki. Laughter exaggerates happy and sad faces depending on visual context. *Psychonomic bulletin & review*, 19(2):163–169, 2012.

[98] William James. The principles of psychology, vol. 1. new york: Henry holt and co, 1890. *EL Thorndike,"Mental Discipline in High-school Studies," Journal of Educational Psychology*, 15:1–22, 1924.

[99] Marc Jeannerod. The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain sciences*, 17(2):187–202, 1994.

[100] Simone Bosbach, Wolfgang Prinz, and Dirk Kerzel. A simon effect with stationary moving stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, 30(1):39, 2004.

[101] Massimo Gangitano, Felix M Mottaghy, and Alvaro Pascual-Leone. Phase-specific modulation of cortical motor output during movement observation. *Neuroreport*, 12(7):1489–1492, 2001.

[102] Marcel Brass, Harold Bekkering, and Wolfgang Prinz. Movement observation affects movement execution in a simple response task. *Acta psychologica*, 106(1-2):3–22, 2001.

[103] Laila Craighero, Arianna Bello, Luciano Fadiga, and Giacomo Rizzolatti. Hand action preparation influences the responses to hand pictures. *Neuropsychologia*, 40(5):492–502, 2002.

[104] Kelly S Wild, Ellen Poliakoff, Andrew Jerrison, and Emma Gowen. The influence of goals on movement kinematics during imitation. *Experimental brain research*, 204(3):353–360, 2010.

[105] Nathan C Foster, Simon J Bennett, Joe Causer, Geoffrey Bird, Matthew Andrew, and Spencer J Hayes. Atypical biological kinematics are represented during observational practice. *Journal of Experimental Psychology: Human Perception and Performance*, 44(6):842, 2018.

[106] Aleksandra Kupferberg, Markus Huber, Bartosz Helfer, Claus Lenz, Alois Knoll, and Stefan Glasauer. Moving just like you: motor interference depends on similar motility of agent and observer. *PloS one*, 7(6):e39637, 2012.

[107] Marco Gandolfo, Vanessa Era, Gaetano Tieri, Lucia Maria Sacheli, and Matteo Candidi. Interactor's body shape does not affect visuo-motor interference effects during motor coordination. *Acta psychologica*, 196:42–50, 2019.

[108] Jung-ran Park. Interpersonal and affective communication in synchronous online discourse. *The Library Quarterly*, 77(2):133–155, 2007.

[109] Karolina Ziembowicz and Andrzej Nowak. Prosody of text communication? how to induce synchronization and coherence in chat conversations. *International Journal of Human–Computer Interaction*, 35(17):1586–1595, 2019.

[110] Steve Whittaker and Brid O'Conaill. The role of vision in face-to-face and mediated communication. 570:23–49, 1997.

[111] Jeremy N Bailenson, Jim Blascovich, Andrew C Beall, and Jack M Loomis. Equilibrium theory revisited: Mutual gaze and personal space in virtual environments. *Presence: Teleoperators & Virtual Environments*, 10(6):583–598, 2001.

[112] T Bickmore and R Picard. Subtle expressivity by relational agents. In *Proceedings of the CHI 2003 Workshop on Subtle Expressivity for Characters and Robots*, 2003.

[113] Sin-Hwa Kang, James H Watt, and Sasi Kanth Ala. Communicators' perceptions of social presence as a function of avatar realism in small display mobile communication devices. In *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008)*, pages 147–147. IEEE, 2008.

[114] Clifford Nass and Youngme Moon. Machines and mindlessness: Social responses to computers. *Journal of social issues*, 56(1):81–103, 2000.

[115] Lee Sproull, Mani Subramani, Sara Kiesler, Janet H Walker, and Keith Waters. When the interface is a face. *Human-computer interaction*, 11(2):97–124, 1996.

[116] Stefan Kopp, Lars Gesellensetter, Nicole C Krämer, and Ipke Wachsmuth. A conversational agent as museum guide–design and evaluation of a real-world application. In *International workshop on intelligent virtual agents*, pages 329–343. Springer, 2005.

[117] Cristina B Gibson. The efficacy advantage: Factors related to the formation of group efficacy 1. *Journal of Applied Social Psychology*, 33(10):2153–2186, 2003.

[118] Sigal G Barsade and Donald E Gibson. Group affect: Its influence on individual and group outcomes. *Current Directions in Psychological Science*, 21(2):119–123, 2012.

[119] Madalina Vlasceanu, Michael J Morais, Ajua Duker, and Alin Coman. The synchronization of collective beliefs: From dyadic interactions to network convergence. *Journal of Experimental Psychology: Applied*, 26(3):453, 2020.

[120] Sigal G Barsade. The ripple effect: Emotional contagion and its influence on group behavior. *Administrative science quarterly*, 47(4):644–675, 2002.

[121] Amy L Baylor. Promoting motivation with virtual agents and avatars: role of visual presence and appearance. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1535):3559–3565, 2009.

[122] Huao Li, Tianwei Ni, Siddharth Agrawal, Dana Hughes, and Katia Sycara. Team synchronization and individual contributions in coop-space fortress. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 64, pages 82–86. SAGE Publications Sage CA: Los Angeles, CA, 2020.

[123] Walter Andrew Schloss. *On the Automatic Transcription of Percussive Music–From Acoustic Signal to High-level Analysis*. Stanford University, 1985.

[124] Stanley Milgram, Leonard Bickman, and Lawrence Berkowitz. Note on the drawing power of crowds of different size. *Journal of personality and social psychology*, 13(2):79, 1969.

[125] Michael J Hove, John R Iversen, Allen Zhang, and Bruno H Repp. Synchronization with competing visual and auditory rhythms: bouncing ball meets metronome. *Psychological Research*, 77(4):388–398, 2013.

[126] Lingyu Gan, Yingyu Huang, Liang Zhou, Cheng Qian, and Xiang Wu. Synchronization to a bouncing ball with a realistic motion trajectory. *Scientific reports*, 5(1):1–9, 2015.

[127] Shuichi Nishio, Tetsuya Watanabe, Kohei Ogawa, and Hiroshi Ishiguro. Body ownership transfer to teleoperated android robot. In *International conference on social robotics*, pages 398–407. Springer, 2012.

[128] Ke Ma and Bernhard Hommel. The role of agency for perceived ownership in the virtual hand illusion. *Consciousness and cognition*, 36:277–288, 2015.

[129] Bigna Lenggenhager, Tej Tadi, Thomas Metzinger, and Olaf Blanke. Video ergo sum: manipulating bodily self-consciousness. *Science*, 317(5841):1096–1099, 2007.

[130] Sotaro Shimada, Kensuke Fukuda, and Kazuo Hiraki. Rubber hand illusion under delayed visual feedback. *PloS one*, 4(7):e6185, 2009.

[131] Sarah-J Blakemore, Chris D Frith, and Daniel M Wolpert. Spatio-temporal prediction modulates the perception of self-produced stimuli. *Journal of cognitive neuroscience*, 11(5):551–559, 1999.

[132] Floris T Van Vugt, Juliane Ritter, Jens D Rollnik, and Eckart Altenmüller. Music-supported motor training after stroke reveals no superiority of synchronization in group therapy. *Frontiers in human neuroscience*, 8:315, 2014.

[133] Adeetee Bhide, Alan Power, and Usha Goswami. A rhythmic musical intervention for poor readers: A comparison of efficacy with a letter-based intervention. *Mind, Brain, and Education*, 7(2):113–123, 2013.

[134] Michelle W Hardy and A Blythe LaGasse. Rhythm, movement, and autism: using rhythmic rehabilitation research as a model for autism. *Frontiers in integrative neuroscience*, 7:19, 2013.

[135] Dana David, Lesly Wade-Woolley, John R Kirby, and Katharine Smithrim. Rhythm and reading development in school-age children: A longitudinal study. *Journal of Research in Reading*, 30(2):169–183, 2007.

[136] Katie Overy. Making music in a group: synchronization and shared experience. *Annals of the New York Academy of Sciences*, 1252(1):65–68, 2012.

[137] D Vaughn Becker, Uriah S Anderson, Chad R Mortensen, Samantha L Neufeld, and Rebecca Neel. The face in the crowd effect unconfounded: happy faces, not angry faces, are more efficiently detected in single-and multiple-target visual search tasks. *Journal of Experimental Psychology: General*, 140(4):637, 2011.

[138] Maya Lecker, Ron Dotsch, Gijsbert Bijlstra, and Hillel Aviezer. Bidirectional contextual influence between faces and bodies in emotion perception. *Emotion*, 20(7):1154, 2020.

[139] Stephane Coˆte. Group emotional intelligence and group performance. In *Affect and groups*, volume 10, pages 309–336. Emerald Group Publishing Limited, 2007.

[140] Sophie van der Woerdt and Pim Haselager. When robots appear to have a mind: The human perception of machine agency and responsibility. *New Ideas in Psychology*, 54:93–100, 2019.

[141] James A Russell. How shall an emotion be called? 484:205–220, 1997.

[142] Paul Ekman and Wallace V Friesen. Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17(2):124, 1971.

LIST OF PUBLICATIONS

Peer reviewed journal papers

- Meneses, A., Mahzoon, H., Yoshikawa, Y., Ishiguro, H. [Article under review] Virtual Agents for Enhancing Emotion Perception and Conveyance".

- Meneses, A., Mahzoon H., Yoshikawa, Y. Ishiguro, H. (2022). Multiple Groups of Agents for Increased Movement Interference and Synchronization. Sensors, 22(14), 5465.

- Meneses, A., Yoshikawa, Y., Ishiguro, H. (2021). Effect of synchronous robot motion on human synchrony and enjoyment perception. Interaction Studies, 22(1), 86-109.

Peer reviewed international conference papers

- Nitada, Y., Yoshikawa, Y., Meneses, A., and Ishiguro, H. (2021, November). Enhancing Sense of Attention from a Communication Robot by Drawing the User's Face on Its Thought Bubble in the Video Conferencing System. In Proceedings of the 9th International Conference on Human-Agent Interaction (pp. 443-447).

Domestic Conference without peer Review

- Meneses A., Yoshikawa Y, Ishiguro H. (2017) Multiple Robot System to Improve Human Behavior Rhythm Synchronization, The 35th Annual Conference of the robotics society of Japan. Interaction (pp. 443-447).

Miscellaneous

- Mita, S., Hatanaka, G., Meneses, A., Thammasan, N., and; Miura, D. (2017, October 23-28). Multi-instrumental end-to-end convolutional neural network for multiple f0 estimation.[Paper presentation] Mirex 2017: In Proc. the 11th Music Information Retrieval Evaluation eXchange, held in conjunction with ISMIR 2017, Suzhou, China.
- Mita, S., Hatanaka, G., Meneses, A., Thammasan, N., and; Miura, D. (2017, October 23-28). Separately training Convolutional Neural Nets for Ensemble category estimation and for Multiple F0 estimation. [Paper presentation] Mirex 2017: In Proc. the 11th Music Information Retrieval Evaluation eXchange, held in conjunction with ISMIR 2017, Suzhou, China.
- Meneses A., Yoshikawa Y, Ishiguro H. (2019, January 23-25) Multiple Robots System for More Synchronizing Human-Robot Interaction [Poster presentation] The 1st International Symposium on Symbiotic Intelligent Systems.
- Meneses A., Yoshikawa Y, Ishiguro H. (2018, January 20-21) Multiple Robot System to Improve Human Behavior Rhythm Synchronization [Poster presentation], 1st International Symposium on Systems Intelligence Division.