

# Evaluation of the effect of mimicry on facial expression in Avatar-Mediated Communication

Mikihiro Suda<sup>1</sup>, Mizuki Oka<sup>1</sup>

<sup>1</sup>University of Tsukuba, Tsukuba, Ibaraki 305-8577 Japan

suda@websci.cs.tsukuba.ac.jp

## Abstract

Against the background of remote work and labor-saving being promoted globally, the use of avatars is becoming widespread in our daily lives. Concurrently, the environments in which avatars are used are also diversifying, with environments appearing wherein communication is possible between humans and avatars as well as between avatars themselves. In this social situation, the effects of use of avatars on communication must be investigated. However, research to compare the effects of non-verbal information in avatar–avatar and human–avatar environments is inadequate. In this study, we created an avatar of which every facial feature can be moved independently and then measured the effects of facial expressions on communication in avatar–avatar and human–avatar environments. The results of a communication experiment based on negotiating in the context of the Prisoner’s Dilemma game showed that mimicking facial expressions resulted in negotiations having a more cooperative outcome. Furthermore, the results suggest that this tendency is stronger in the avatar–avatar environment.

## Introduction

Covid-19 pandemic affected human communication significantly. As face-to-face communication was discouraged, video chatting via the Internet such as, the Zoom video conferencing and Microsoft Teams applications, became increasingly used, and tools such as FaceRig<sup>1</sup> and Snap Camera<sup>2</sup>, which introduce avatars into video chatting, attracted greater attention. These tools use web cameras to analyze facial positioning and user’s facial rotation while making expressions, and then locate and introduce virtual objects (avatars) into the video data. Among them, the use of FaceRig is growing remarkably, with the total number of users in 2020 being 275% larger than in 2018<sup>3</sup>. In part, this growth in the use of avatars is due to the social ubiquity, lower cost, and higher performance of personal computers (Reeves, 2012). The expected trend following the COVID-19 epidemic is that the increasingly widespread use of computers

will continue, along with the concomitant spread of avatar use.

The process of creating, exchanging, and perceiving information by encoding, transmitting, and decoding messages using a computer is called computer-mediated communication (CMC) (Jonassen and Driscoll, 2003). Avatar-mediated communication (AMC) is a typical CMC because it uses computers to display avatars and transmit various information (L. and Fox, 2018). In the context of CMC, researchers have studied the types and number of information exchanged and the attributes of the people engaging in CMC. Non-verbal expressions exchanged in communication are known as social cues. Social cues such as gesture, gaze, head movement, and breathing are fully exchanged in face-to-face communication; however, in CMC, social cues are partially lost because the communication is processed by the computer. Based on the loss of social cues, many studies have reported on the inferiority of CMC compared to face-to-face communication (Walther and Parks, 2002). For example, Rutter and Stephenson (1979) reported that the number of social cues that can be exchanged has a significant effect on communication (cuelessness model). These studies reporting the recessive nature of CMC based on the loss of social cues have been summarized as the cue filtered-out model, and it is now widely known that deficits in social cues by computer negatively affect communication.

However, as multimedia develops, some argue that the inferiority of CMC is decreasing. With multimedia usage spreading, a concept called social presence has been attracting attention. Biocca (1997) defined social presence as “the subjective feeling of being with a real person and being able to touch his or her thoughts and feelings”. In their review paper on social presence, Oh et al. (2018) reported that many studies suggest that multimedia enhances social presence compared to text-only communication. In other words, multimedia brings CMC closer to face-to-face communication.

In addition to multimedia, avatar-mediated communication (AMC) is thought to have the same effect to eliminate the recessive nature of CMC (Bente et al., 2008). This is because AMC allows more social cues than a traditional

<sup>1</sup><https://facerig.com/>

<sup>2</sup><https://snapcamera.snapchat.com/>

<sup>3</sup><https://steamdb.info/app/274920/> (Retrieved on 2021-01-10)

medium such as text or sound as avatars can reproduce and transmit non-verbal expressions such as facial expressions and gestures. To understand the characteristics of AMC, we need to understand those of face-to-face communication.

Before the CMC study, face-to-face communication was the subject of many psychological studies and numerous reports have been made thereon. As mentioned above, non-verbal expressions exchanged in face-to-face communication are also known as social cues, and include gestures, breathing, facial expressions, head movements, etc. Pentland (2010) especially focused on non-verbal expressions which are exchanged unconsciously, which he called honest signals. He argues that we can accurately predict the outcome of communication by analyzing face-to-face communication of short duration based on the four components of honest signals: mimicry, activity-level, cooperation, and agitation. For example, Curhan and Pentland (2007) shows that a task requiring high-level cooperation such as salary negotiations with one's boss demands high cooperation, high mimicry, and low agitation. Or, Stoltzman (2006) shows that a task requiring high-level leadership such as selling a new business plan calls for high activity-level, high cooperation, and low agitation.

Mimicry, which Pentland (2010) refers to as one of the honest signals, is the phenomenon of humans unconsciously imitating the words and actions of others in communication. He argues that mimicry has the effect of increasing positive feelings and trust. It is known that some animals including humans instinctively mimic others, and some researchers like Masumori et al. (2021) focus on the learning effect of mimicry.

The effect of a non-verbal expression like mimicry is one of the targets of avatar-mediated communication (AMC) studies. Especially, mimicry is one of the most focused study targets and many studies suggest that mimicry has some effect on communication despite it being unconsciously communicated to participants (Hale and Hamilton, 2016; Bailenson and Yee, 2005). Bogdanovych et al. (2018) conducted an experiment based on the Prisoner's Dilemma game using an avatar and mimicry and found that mimicry increases the smile quotient of communication participants. According to such studies, we can evaluate that non-verbal expression such as the honest signals affect face-to-face communication and also (perhaps partially) avatar-mediated communication.

Thus, the investigation of non-verbal expressions in avatar-mediated communication has progressed on aspects of both CMC and face-to-face communication studies. However, previous avatar-mediated communication studies have been conducted only in human-avatar environments or avatar-avatar environments and there have been no cross-sectional comparative studies of the two combined. Despite avatars having become diversified in recent years, and now demonstrate high performance and high accuracy, there is

a lack of research on the effects of avatar performance and accuracy on communication.

The most important characteristic of the avatar-avatar environment is that the self is represented as an avatar. Moreover, as the user's face and the communicating person's face were juxtaposed in the video chatting interfaces, the users tended to view their own face during the communication. In this environment, it is important to investigate and understand the sense of ownership; the sense in which the avatar is a part of myself, and the sense of agency; the sense in which the communication outcome results from my actions. In this study, we investigate how mimicry affects communication in both the avatar-avatar environment and human-avatar environment, which mimics the video chatting systems. Does the mimicry effect arise in the avatar-avatar environment and human-avatar environment similar to the face-to-face environment? In this study, we conducted an experiment of two-party communication based on free-talk negotiation using the Prisoner's Dilemma scenario.

## Avatar

### Existing Avatars

Avatars on the open market have become highly diverse in recent years. In this section, we give an overview of these avatars in Table 1. We categorize them as static or dynamic, then the dynamic avatars are further classified by their realism and operability.

Static avatars are of the type for which users create an illustration of themselves and send or receive it as an image or a short video via chat or messaging system. Facebook Avatar by Facebook <sup>4</sup>, LINE Avatar by LINE <sup>5</sup>, Bitmoji by Snap <sup>6</sup> are classified as static avatars. Using these tools, users can create avatars by combining prepared illustrations of body parts. Of static avatars, it can be said that they are Emoji variants since they are already-created images.

Dynamic avatars are the type that follows the user's behavior in real-time and can be processed as a streaming video. Memoji by Apple <sup>7</sup>, MetaHuman by Epic Games <sup>8</sup>, open-source software MakeHuman <sup>9</sup>, Japanese platform Live2D <sup>10</sup> and Reality <sup>11</sup> are classified as dynamic avatars. Dynamic avatars are further classified by their life-like appearance. Memoji, Live2D, and Reality provide anime-like avatars. Live2D, especially, is a tool for creating dynamic avatars from 2D illustrations and has high acceptance in Japanese pop culture. Conversely, MakeHuman and

<sup>4</sup><https://www.facebook.com/help/278747370042382>

<sup>5</sup><https://linecorp.com/ja/pr/news/ja/2020/3427>

<sup>6</sup><https://www.bitmoji.com/>

<sup>7</sup><https://support.apple.com/en-us/HT208986>

<sup>8</sup><https://www.unrealengine.com/en-US/digital-humans>

<sup>9</sup><http://www.makehumancommunity.org/>

<sup>10</sup><https://www.live2d.com/>

<sup>11</sup><https://reality.app/>

Table 1: Comparison of avatars on the market

	device	part		operability		appearance		exportability
		face	body	static	dynamic	anime	real	
MakeHuman	any	✓	✓		✓		✓	✓
Bitmoji	Android, iOS	✓	✓	✓		✓		
Meta Humans	any	✓			✓		✓	✓
Live2D	Windows, Mac	✓	✓		✓	✓		✓
Reality	Android, iOS	✓	✓		✓	✓		✓
Memoji	iOS	✓			✓	✓		
Facebook Avatar	Android, iOS	✓	✓	✓		✓		
LINE Avatar	Android, iOS	✓		✓		✓		

MetaHuman provide realistic avatars that are used in games or academic research.

Moreover, dynamic avatars can also be classified by their operability. On Live2D or Reality, users manually change an avatar’s pose or expression. Conversely, Memoji uses the camera of an iOS device to capture varied views of the user’s facial features and can reproduce rich facial expressions by coordinating them with the avatar’s moving parts. The MakeHuman and MetaHuman avatars created with their higher degrees of free-operation, are exported as generic 3D models and can be freely used by developers on 3D engines such as Unity or Unreal Engine. These avatars have movable body parts, thus developers can reproduce various kinds of poses and expressions by controlling the mobile parts such as the elbow joint or neck axis.

As described above, the avatars in the market are becoming increasingly diverse, but, for the reasons below, they are difficult to use directly for academic experiments. First, most avatars cannot be used on open development platforms because they can only be used on their platforms. For instance, although Memoji by Apple has high-performance replication of facial expressions the Memoji functions are not accessible. Second, the avatar’s standard of movement is low. The movements of large parts, such as arms and legs, can, to some extent, be developed in a generalized manner, but relatively detailed movements such as facial expression movements vary depending on the platform and avatar.

## Our Avatar Model

For this reason, we create a new avatar for our experiments. The avatar is shown in Fig. 1. It is based on a 3D model distributed on TurboSquid<sup>12</sup>, 3D model share site, and facial expression animations were added to the model with Blender 2.83<sup>13</sup>. The demonstration of our avatar is available on the web<sup>14</sup>.

<sup>12</sup><https://www.turbosquid.com/3d-models/male-head-obj/346686>

<sup>13</sup><https://www.blender.org/>

<sup>14</sup>[https://sudame.github.io/avatar\\_web/](https://sudame.github.io/avatar_web/)

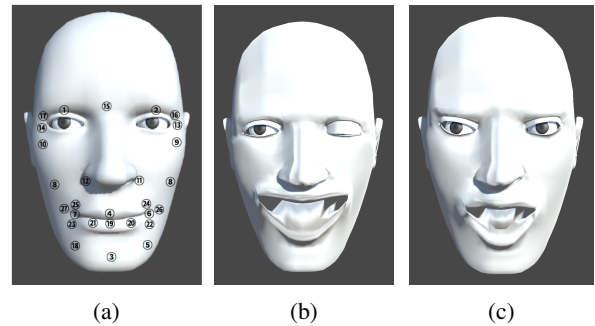


Figure 1: The avatar used in the experiment. The avatar can manipulate 27 points on the face and the eyes independently and can generate various facial expressions. (a) The natural face. The approximate locations of the moving parts are numbered. (b, c) The laughing face and the angry face. The avatar can generate rich facial expressions.

Our avatar works with Apple ARKit 4<sup>15</sup>, Apple’s augmented reality library. Using ARKit, we can capture detailed movements of many of the device user’s facial parts through the iOS device camera. Our avatar generates rich facial expressions flexibly by synchronizing the avatar’s movable parts and the movement data of the user’s facial parts collected using ARKit. ARKit can collect 52 parts on a face and analyze their movements, but these include very small movements and it is difficult to reproduce them all. Thus we selected only 27 details that produce relatively large movements such as eyes, mouths, and cheeks, and applied them to the avatar. In addition to facial movements, our avatar can rotate its eyeballs. We synchronized the avatar’s eyeballs and the avatar users’ gaze collected through ARKit, so our avatar can reproduce very natural eye movements. Our avatar can move each detail — the facial parts and each eyeball — independently. In the experiments described below, this function was used to attain accurate mimicry.

<sup>15</sup><https://developer.apple.com/augmented-reality/arkit/>

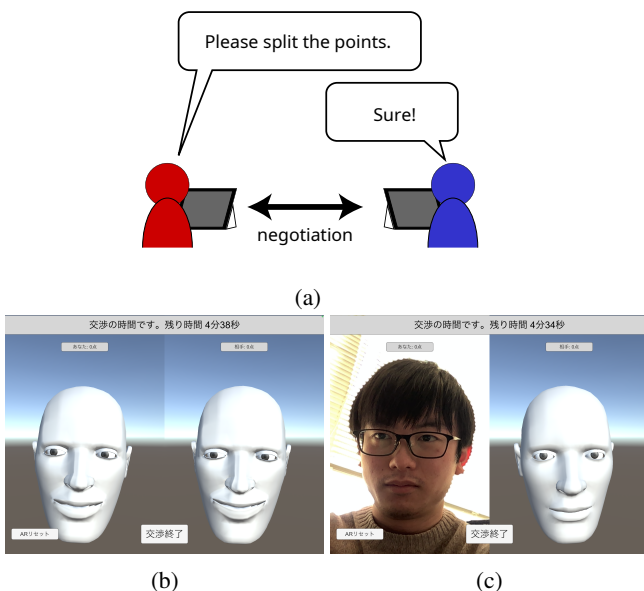


Figure 2: Experimental environment. (a) The participants and the experiment operator negotiate using a tablet device. (b, c) The tablet’s display. It shows the self (a participant) on the left half and the opponent (an operator) on the right. (b) Avatar–avatar environment. A participant and an operator using the same avatars. (c) Human–avatar environment. Self (a participant) is displayed as live video on the left and the opponent (an operator) is displayed as an avatar on the right.

## Experiments

With the cooperation of 16 participants, we experimented to compare the effects of facial expression mimicry in the avatar–avatar environment and the human–avatar environment. In this experiment, a participant and an experiment operator conducted negotiations based on the Prisoner’s Dilemma. We used tablet devices to negotiate and switched the avatar–avatar environment and the human–avatar environment by changing the participant’s display.

As described in the introduction, the mimicry effect increases positive feelings and trust. Thus it is expected that under the mimicry condition the negotiation will conclude with the cooperative result. In addition to this, it is expected that the presence and scale of the mimicry effect will change between the avatar–avatar environment and the human–avatar environment.

## Environment

Here, we created an application having an interface similar to common video chatting applications, such as Zoom or Microsoft Teams. Fig. 2a shows the experimental setup. The negotiation is conducted via Apple tablets, which are placed in front of each participant. The iOS application prepared for the experiment uses the front camera of the tablet

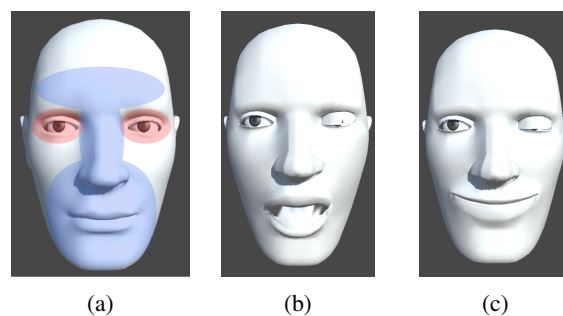


Figure 3: The avatar’s mimicry settings. (a) Areas to be mimicked. Red areas are mimicked, blue areas are not. (b, c) (c) is mimicking (b). The eye motions are mimicked, so their movements match, but the movements of other parts do not match.

Table 2: Four experimental conditions

	mimicry	not mimicry
avatar – avatar	condition 1	condition 2
avatar – human	condition 3	condition 4

and then reproduces the participant’s behavior on an avatar on the tablet. For the task environment, we prepared the avatar–avatar environment (Fig. 2b) and human–avatar environment (Fig. 2c). In the avatar–avatar environment, both participants (experiment participant and experiment operator) are displayed as avatars and in the human–avatar environment the experiment operator is displayed as an avatar but the experiment participant is displayed as a live video. Because we prepared only one kind of avatar, in the avatar–avatar environment each participant used the same avatar.

## Mimicry

We conducted experiments with and without mimicry in the avatar–avatar environment and the human–avatar environment. Therefore, a total of four conditions were employed (Table 2). Fig. 3 shows the setting of mimicry in the avatar. Under the mimicry condition, an avatar of the experiment operator — the negotiation opponent — mimicked a participant’s facial expression. In this experiment, the opponent did not mimic the entire facial expression but only the eye movements and the opening and closing of the eyelids. This means that the red area in Fig.3a was mimicked (imitated the participant’s movement) but the blue area was not mimicked (only reproducing the opponent’s original movement) (Fig. 3b, 3c). The purpose of these settings was to prevent the mimicry from being exposed due to the discrepancies between the speech and the mouth movements. For the same reason, we delayed the mimicry for two seconds, differently expressed, the opponent avatar reproduced move-

ments of the participant's eyeballs and eyelids with a two seconds delay.

### Task

In this experiment, we conducted a negotiation task called *the Prisoner's Dilemma*. This task is a game in which, within five minutes, two parties negotiate for points and take points according to the result of the negotiation (Burton-Chellew and West, 2012). Participants can choose *Steal* or *Split* as the next action after negotiation. If both participants choose *Split*, each takes half the points. However if one chooses *Steal* and the other *Split*, the one who chose *Steal* takes all the points, while the other takes none. If both choose *Steal*, neither takes a point. In this game, when the participant wants more points they need to choose *Steal* but if each participant thinks the same, neither of them takes any points and this is a so-called dilemma. Negotiation has an important role in this game to avoid one's risk and maximize one's point. Participants have negotiation and action choices three times in each game, and every game they negotiate for 100 points. Theoretically, this means they can earn a maximum of 300 points and they work together on the task to earn higher points.

### Experimental procedure

On arrival at the experimental room, we explained the experiment and the experimental data handling to the participants and asked them to sign the experimental consent form. When we explain the experiment, we avoid telling them that the avatars mimic their facial expressions. After the explanation, a tablet fixed with a tablet stand was placed on the desk in front of the participant and the participant was asked to adjust the angle of view and the position of the device by starting the camera application. Then we launched the experimental application and explained the task in the avatar–avatar environment without mimicry. Then, we explained that the tablet shows the participant's avatar on the left and the opponent's (the experiment operator's) avatar on the right, and asked the participant for confirmation. To check the synchronization of the movements between the participant's own head and the avatar, we explained that the movements of mouth, cheek, eyes, nose, and rotation of the head were synchronized with the avatar, and asked the participant to move these parts while looking at the avatar for about one minute. After that, we explained the rules of the negotiation task and asked the participant to operate the experimental application as a test. In the test, we asked them to choose *Steal* and the experiment operator chose *Split*.

During the task, participants could freely talk to the opponent. To encourage active negotiation, the experiment operator asks the participant questions such as "How honest are you going to be this time?" or "What is your strategy this time?" To minimize the effect of the operator's action choices, the operator's choices were automatically created

based on the rules. This was a within-subject experiment. The subjects conducted the experiments in all the four conditions. As the condition-setting order was expected to affect the experimental outcome, the order was randomized for each subject. Consequently, the experimental operator responses, and their combination with the conditions were randomized. This was hypothesized to lower the influence on the experimental bias. As soon as all the tasks were completed the operator asked the participant to fill out a questionnaire.

### Evaluation

We evaluated the experiment from two aspects; objective and subjective evaluation methods.

We use the action selection data from each condition and the facial expression data acquired to control the avatar for the objective evaluation. The action selection data is binary data that has *Steal* and *Split*. The subjects selected these actions thrice per game and each subject conducted games under the four conditions, thus we obtained twelve action selection data items from each participant. The facial expression data is the data we obtained to control the avatar. We recorded the facial expression data every three seconds and used the data for the analysis.

For the objective evaluation we used the God Speed method, which is widely found in the field of HRI (Human-Robot Interaction) as an index of human subjective impressions of robots in human-robot interactions, human-agent interactions, or human-avatar interactions (Bartneck et al., 2009). The God Speed method questionnaire consists of 23 questions using the SD (Semantic Differential) method and we can analyze the impression from five perspectives: anthropomorphism, likeability, perceived intelligence, perceived safety, and animacy.<sup>16</sup>

### Result

The 16 healthy adult male and female participants were recruited by word of mouth in December 2020. One participant had to stop due to a problem in the application for the experiment that affected its execution, so we excluded data related to that participant from the analysis. Besides this, we could not record accurate facial expression data of three of the participants due to a logging system error. As a result, the data of 15 participants were used for the analysis of action selection data, and the answers to the questionnaire. The data of 12 participants were used for analyzing the facial expression data.

<sup>16</sup> Animacy is the sense of an artifact being alive. As reported in the "uncanny valley phenomenon", it is known that the sense of being alive does not necessarily correlate with the target object being realistic

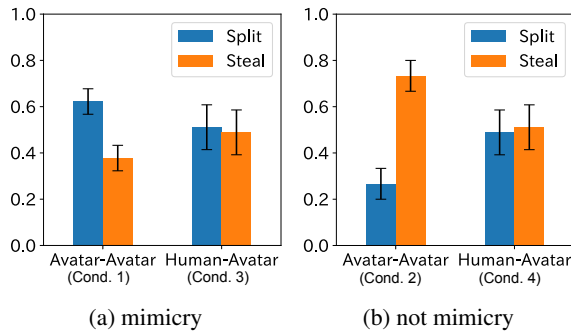


Figure 4: Results of action selection in a task. When mimicry is performed (cond.1 and cond.3), the rate of selection for *Split* increases and when it is not performed (cond.2 and cond.4), the rate of selection for *Steal* increases. Both tendencies are more pronounced in the avatar–avatar environment (cond.1 and cond.2).

### Action Selection

Fig. 4 shows the result of the action selection data. In the mimicry condition (cond.1 and cond.3), *Split* was selected 57.3% of the time (Fig. 4a) and in the non-mimicry condition (cond.2 and cond.4), *Split* was selected 36.3% of the time (Fig. 4b). We conducted a hypothesis test for the population proportions (significant level 1%) and the result shows that there is a significant difference in action selection rate between mimicry and non-mimicry conditions ( $\chi^2 = 7.79$ ,  $p = 0.00525$ ). This indicates that mimicry has a significant effect on increasing the split-selection rate.

Moreover, this tendency is stronger in the avatar–avatar condition (cond.1 and cond.2). However, because of insufficient samples, a hypothesis test for the population proportions (significant level 5%) did not show significant difference(mimicry condition (cond.1):  $\chi^2 = 0.900$ ,  $p = 0.343$ , non-mimicry condition (cond.2):  $\chi = 3.742$ ,  $p = 0.053$ ).

### Facial Expression

We compared the rate of facial expression variation between environments to see how they are related. It is not appropriate to compare directly the variation of facial parts between participants because of individual differences in the magnitude of facial expressions. In addition, some parts such as the mouth and eyes display relatively large movements and some parts such as the nose display relatively small movements. Consequently, it is not appropriate to compare directly the movements between facial parts to understand the difference of facial expressions by movement size. Therefore, we used a relative score, the rate of facial expression variation, for analysis. The score is based on each participant’s mean facial expression variation in the human–avatar environment and under the non-mimicry condition (cond.4). Also, eight parts were excluded for which, from the analysis, the movement size was zero for more than half of the

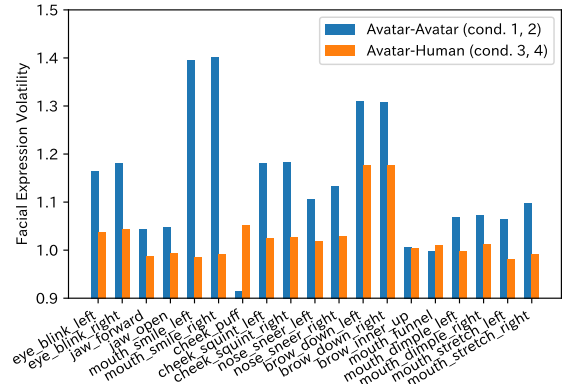


Figure 5: A comparison of the rate of facial expression variability in the avatar–avatar and the human–avatar environments (cond.1,2 vs cond.3,4). For most of the facial expressions the score is higher in the avatar–avatar environment.

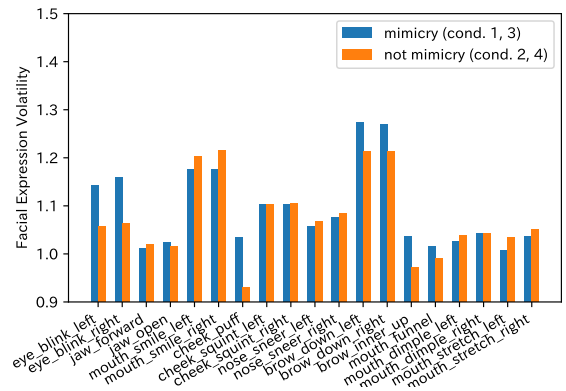


Figure 6: A comparison of the rate of facial expression variability in the mimicry and the not-mimicry conditions (cond.1,3 vs cond.2,4). The difference is not significant between two conditions.

data due to small facial movements and the limited technical measurement.

Fig. 5 shows the result of comparing the variability in facial expression between the avatar–avatar environment and the avatar-human environment (cond.1,2 vs cond.3,4). We conducted one-way analysis of variance (ANOVA) and the result shows there is a significant difference between the rate of facial expression variation between the avatar–avatar environment and the human–avatar environment( $F = 18.0$ ,  $p = 0.000$ ). This means that the avatar–avatar environment has a significant effect on expanding the variation in facial expressions.

Fig. 6 depicts the facial movement differences between the mimicry and not-mimicry conditions. We did not observe significant differences ( $F = 0.382$ ,  $p = 0.538$ ) upon employing the one-way analysis of variance (ANOVA).

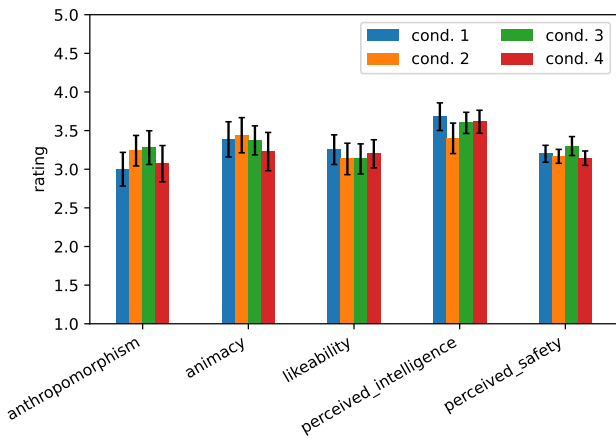


Figure 7: Results of the God Speed questionnaire. No significant differences were observed in any of the evaluation aspects. The error bars indicate standard error of mean.

### God Speed Questionnaire

We analyzed answers of the questionnaire using one-way ANOVA for each of the aspects of the God Speed method: anthropomorphism, animacy, likeability, perceived-intelligence, and perceived-safety. The result is depicted in Fig. 7 and it shows that there is no significant difference between each experimental conditions (cond.1, cond.2, cond.3, and cond.4) (anthropomorphism:  $F = 0.383$ ,  $p = 0.765$ , animacy:  $F = 0.159$ ,  $p = 0.923$ , likeability:  $F = 0.0914$ ,  $p = 0.964$ , perceived-intelligence:  $F = 0.527$ ,  $p = 0.666$ , perceived-safety:  $F = 0.433$ ,  $p = 0.731$ ). This means that avatar-using environment and mimicry does not have any effect on subjective impressions measurable by the God Speed method.

We also analyzed answers to the post-experiment questionnaire that asked participants whether they perceived mimicry or not and found that every participant didn't perceive mimicry.

### Discussion

In this study, we investigate the effect of mimicry on action selection in the negotiation task and we found that mimicry effects increase the rate of choosing more cooperative action. This result supports existing research and suggests that mimicry affects impressions such as increasing positive feelings and trust. Moreover, we found that mimicry in the avatar–avatar environment may be more effective than in the human–avatar environment.

Here, we demonstrate that the subjects exhibited greater facial movements in the avatar–avatar environment than in the avatar–human environment. However, significant facial movement differences were not observed between the mimicry and not-mimicry conditions. Furthermore, stronger mimicking effect in the avatar–avatar environment sug-

gested a relationship between the mimicking effect and the subject facial movement magnitude. Hence, larger facial movements might increase the mimicking effect.

This study is different from the previous studies, as the participants here communicated while looking at their faces, which potentially affected the outcome of this study. The non-verbal expression differences between the communications conducted a) while looking only at the other person's face and b) looking at their own face, can be investigated in the future. Additionally, in the avatar–avatar environment, the two participants used the same avatar to communicate. That both communicators used the same avatar may be a reason for the increased effectiveness of mimicry in the avatar–avatar environment. Most of the previous studies on an avatar's appearance have focused on the appearance of the communicating partner, and few studies have focused on its own appearance. Conversely, with the widespread use of video chats and other forms of communication, the avatar–avatar environment is being more commonly used for communicating. In the future, there is a need to investigate the effect of the relationship between one's own appearance and that of one's interlocutor during communication.

The stronger effect of mimicry in the avatar–avatar environment may be explained by the hyper-personal theory, which argues that computer-mediated communication is more effective than face-to-face communication in certain tasks because unnecessary social cues are removed by computers (Walther, 1996; Walther and Whitty, 2021). In the human–avatar environment, participants are required to process many kinds of information unrelated to mimicry such as a hairstyle or the background because participants were shown as live-action videos. Conversely, in the avatar–avatar environment, such unrelated information was removed and participants could pay more attention to facial expressions related to mimicry. Furthermore, because it is expected that participants would pay more attention to the opponent's facial expressions this increases the effect of mimicry. Based on this theory, it is suggested that the avatar–avatar environment is relatively suitable for a task that requires careful reading of the opponent's facial expression such as the negotiation task we conducted.

The subjective evaluation, using the God Speed method, of communication between the conditions did not show the difference, even though the objective evaluation showed that the avatar–avatar environment increased participants' facial expressions and the performance of mimicry increased the proportion of participants who chose a cooperative action. In the analysis of facial expression data, it is noteworthy that there was no change in subjective evaluation, even though the variability of `mouth_smile_left` and `mouth_smile_right`, the region that produces a smiley mouth, is increased by about 40% in the avatar–avatar environment. However, based on the honest signals approach, it is natural that non-verbal expressions affect communica-

tion unconsciously. In the book, “Honest Signals”, Pentland (2010) argues that in different situations humans can make decisions based only on thin slices of observational data yet successfully predict subsequent behavior. According to this argument, regardless of participants’ conscious (subjective) impressions in the negotiation task conducted in this study, the outcome of the negotiation can be predicted to some extent through the implementation of mimicry and how avatars are used. It means that the avatars’ facial expressions functioned as an honest signal and by controlling them we were able to unconsciously influence the outcome of the negotiation.

## Conclusion

In this paper, we discussed avatars, which are increasingly used in the real world and have shown that they are becoming more diversified with higher performance. In addition, we indicated the lack of research into comparing the effects of non-verbal information in avatar–avatar and human–avatar environments and experimented with our self-created avatar. This avatar generates facial expressions by moving the mobile facial regions independently.

We experimented to confirm the effect of mimicking facial expressions in both avatar–avatar and human–avatar environments. The results of the experiment showed that mimicry led to more cooperative negotiation results in both the avatar–avatar and human–avatar environments. Furthermore, this tendency was observed to be stronger in the avatar–avatar environment. Conversely, there was no change in the subjective impressions from the questionnaire survey. It is interesting to note that while no change in the subjective impressions was reported, there was a change in the behavior.

It was suggested that the degree of cooperation in negotiation increased only by having avatars mimic mechanically and we discussed this suggestion based on the sense of ownership. In the future, we will conduct further experiments using diverse avatars and environments to investigate how the various attributes of avatars and environments affect the human senses. What is the relationship between accuracy and mimicry? Is the degree of realism associated with gaining people’s cooperation? As the use of avatars expands in society, we believe it is important to focus on these questions.

In life, sense has an important role in fostering cooperation with others. Through the repeated observation and understanding of senses in virtual space, we may be able to generate senses in the future. We expect that understanding and the generation of senses will, from a new perspective, make a significant contribution to artificial life research.

## References

Bailenson, J. N. and Yee, N. (2005). Digital Chameleons: Automatic Assimilation of Nonverbal Gestures in Immersive Vir-

- tual Environments. *Psychological Science*, 16(10):814–819.
- Bartneck, C., Kulić, D., Croft, E., and Zoghbi, S. (2009). Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots. *International Journal of Social Robotics*, 1(1):71–81.
- Bente, G., Rüggenberg, S., Krämer, N. C., and Eschenburg, F. (2008). Avatar - Mediated Networking: Increasing Social Presence and Interpersonal Trust in Net - Based Collaborations. *Human Communication Research*, 34(2):287–318.
- Biocca, F. (1997). The Cyborg’s Dilemma: Progressive Embodiment in Virtual Environments. *Journal of Computer-Mediated Communication*, 3(2):0–0. JCMC324.
- Bogdanovych, A., Richards, D., Simoff, S., Pelachaud, C., Heylen, D., Trescak, T., Hoegen, R., Schalk, J. v. d., Lucas, G., and Gratch, J. (2018). The impact of agent facial mimicry on social behavior in a prisoner’s dilemma. *Proceedings of the 18th International Conference on Intelligent Virtual Agents*, pages 275–280.
- Burton-Chellew, M. N. and West, S. A. (2012). Correlates of Cooperation in a One-Shot High-Stakes Televised Prisoners’ Dilemma. *PLoS ONE*, 7(4):e33344.
- Curhan, J. R. and Pentland, A. (2007). Thin slices of negotiation: Predicting outcomes from conversational dynamics within the first 5 minutes. *Journal of Applied Psychology*, 92(3):802–811.
- Hale, J. and Hamilton, A. F. D. C. (2016). Testing the relationship between mimicry, trust and rapport in virtual reality conversations. *Scientific Reports*, 6(1):35295.
- Jonassen, D. and Driscoll, M. (2003). Handbook of Research on Educational Communications and Technology.
- L., Nowak, K. and Fox, J. (2018). Avatars and computer-mediated communication: a review of the definitions, uses, and effects of digital representations. *Review of Communication Research*, 6:30–53.
- Masumori, A., Maruyama, N., and Ikegami, T. (2021). Personogenesis Through Imitating Human Behavior in a Humanoid Robot “Alter3”. *Frontiers in Robotics and AI*, 7:532375.
- Oh, C. S., Bailenson, J. N., and Welch, G. F. (2018). A Systematic Review of Social Presence: Definition, Antecedents, and Implications. *Frontiers in Robotics and AI*, 5:114.
- Pentland, A. (2010). *Honest Signals: How They Shape Our World*. The MIT Press. MIT Press.
- Reeves, S. (2012). Envisioning Ubiquitous Computing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI ’12, page 1573–1582, New York, NY, USA. Association for Computing Machinery.
- Rutter, D. R. and Stephenson, G. M. (1979). The Role of Visual Communication in Social Interaction. *Current Anthropology*, 20(1):124–125.
- Stoltzman, W. T. (2006). Toward a social signaling framework : activity and emphasis in speech. *Thesis (M. Eng.)—Massachusetts Institute of Technology, Dept. of Electrical Engineering and Computer Science*.



- Walther, J. B. (1996). Computer-Mediated Communication. *Communication Research*, 23(1):3–43.
- Walther, J. B. and Parks, M. R. (2002). Cues Filtered Out, Cues Filtered in: Computer Mediated Communication and Relationships. 3:529—563.
- Walther, J. B. and Whitty, M. T. (2021). Language, Psychology, and New New Media: The Hyperpersonal Model of Mediated Communication at Twenty-Five Years. *Journal of Language and Social Psychology*, 40(1):120–135.