Computers in Human Behavior Reports 8 (2022) 100239

Contents lists available at ScienceDirect

Computers in Human Behavior Reports

journal homepage: www.sciencedirect.com/journal/computers-in-human-behavior-reports

Study of the characteristics of ear animations used to convey information and emotion in remote communication without web camera

Toshiaki Kakii [a,*], Hideyuki Fujiu [b], Guiming Dai [a]

[a] Sumitomo Electric Industries Ltd., Osaka, Japan
[b] Department of Psychology, Faculty of Human Sciences, University of Tsukuba, Tokyo, Japan

A R T I C L E I N F O

Keywords:
Remote communication
Nonverbal communication
Avatars
Animations
Videoconferencing

A B S T R A C T

The use of remote communication has grown globally due to the COVID-19 outbreak. In some remote communication, meeting participants use audio only with their web cameras turned off, resulting in a lack of exchange of nonverbal information. In this study, we defined an "ear animation" as an avatar composed of a simple face-like body with no facial features and ear-like parts coming out from this body which can be animated. The purpose of this study was to design the ear animation and evaluate user impressions of it as nonverbal information. While setting conveying information and conveying emotion as dependent variables, the independent variables we set in this study were three different conditions: when ear animations were presented with no sound, when ear animations were presented simultaneously with simple voice, when only voice was played, and three different kinds of content: "agreement", "skepticism", and "disagreement" conveyed from ear animations. Using Two-way ANOVA (repeated) with these variables, we conducted comparative analysis. The results showed that ear animations presented simultaneously with voice had the potential to be a new way of conveying nonverbal information by combining relevant ear animation movement forms.

1. Introduction

Due to the effects of COVID-19, the frequency of remote communication is increasing rapidly. For example, it has been reported that Microsoft Teams marked 2.7 billion meetings in one day in March 2020 (Spataro, 2020). Good quality video has been made possible by the widespread availability of high-speed internet and high-performance PCs and other devices. While dynamic picture images such as facial expressions can be transmitted, there are occasions where participants join remote meetings with their cameras turned off. In a survey that asked why you do not show your face in remote communication (Castelli & Sarvary, 2021; Tobi et al., 2021), the most frequent reasons given were: "other people turning their video off", "people wanting to multitask", "people feeling self-conscious about their appearance", "not actively participating" and "the effort of being seen". In short, "feelings of not wanting to be seen", which has nothing to do with the technological aspect of image transmission, has become an obstacle to transmission of facial information. In remote communication where only voice and documents are shown as faces are hidden, nonverbal communication occurring with dialogue is not conveyed enough. Nonverbal information is an important factor for smooth interaction (Archer & Akert, 1997; Haase & Tepper, 1972). Therefore, remote communication with faces not showing on screen prevents the establishment of smooth communication (Rodeghero et al., 2021). Giving feedback and showing reaction from dialogue are especially important forms of nonverbal communication. However, current remote communication has a problem with nonverbal communication not being conveyed effectively (Peper et al., 2021). Communication where no facial expressions appear lacks this reaction to dialogue, so that big problems are created such as a reduction in the sense of liveliness in communication, leading to physical fatigue.

Many studies have been done on the use of avatars as a new method for remote communication where no real faces are shown (Bailenson, 2018; Nowak & Fox, 2018). The results of comparing a virtual reality meeting with a video conferencing environment revealed that avatars improved feelings of presence, closeness, and arousal in the virtual reality environments (Campbell et al., 2019). However, using avatars also causes issues. Especially in business videoconferencing, avatars can be both an enabler and an obstacle in helping interaction with other participants (Forsberg & Kirchner, 2021; Junuzovic et al., 2012). In designing avatars, it is necessary to examine whether an avatar's appearance can effectively interact with user's expression. In order to do

* Corresponding author.
E-mail address: kakii-toshiaki@sei.co.jp (T. Kakii).

https://doi.org/10.1016/j.chbr.2022.100239
Received 23 November 2021; Received in revised form 22 September 2022; Accepted 3 October 2022
Available online 13 October 2022

so, we need to examine how the avatar works (Oh et al., 2016). A study regarding avatars used in video games examines how customization and identification of avatars influence users' communication behaviors (Takano & Taka, 2022). If we make the avatar's face too cartoonish, it becomes too entertaining to be appropriate for business use. Some avatar designs could be misleading in business meetings. This is to say that while avatars have big potential, there are also design-dependent difficulties for individual avatars. The impact of subtle avatar facial design must be considered when avatars are introduced in videoconferencing. The uncanny valley problem, which decreases familiarity once one's avatar resembles to oneself too closely, has been pointed out as an example of this (Mori, MacDorman, & Kageki, 2012; Shin et al., 2019). These issues cannot be solved simply by technical development. In order to deal with situations where a web camera is turned off, it may be beneficial to investigate conveying nonverbal information by means other than avatars equipped with facial expressions, since it will create more options available for diverse communications.

Given these circumstances, we became interested in creating a new interface by asking ourselves a question: "Is it possible to establish lively communication with a sense of unity and presence in a remote communication by using nonverbal communication unrelated to facial expressions?" We thought if simple animations are utilized as nonverbal communication, such an interface would be available for a wide range of people with easy and minimum customization.

In real-time video interaction through monitors, there is an issue with the discomfort in maintaining direct eye contact. Countermeasures for this have been discussed in some studies (Bohannon et al., 2013; Park et al., 2021). Our proposed interface is also expected to be able to end this uncomfortable feeling by eliminating direct eye contact.

2. Purpose and structure of this study

An avatar is defined as a character representing a particular person on the internet. In this study, an avatar composed of a simple face-like body with no facial features and ear-like parts coming out from the body which can be animated was defined as an "ear animation".

The purpose of this study was to design the ear animation and evaluate user impressions of it as nonverbal information. This study consists of three parts. First, we designed movement forms as ear animations for nonverbal communication, referring to human body movements such as head movements and gestures. Next, we conducted an experiment to rate impressions of the ear animations' movement forms. Then we compared psychological impressions using Two-way ANOVA (repeated) with independent variables. As independent variables, we set three different types of content conveyed by the ear animations and three different conditions. The three types of content conveyed were: "agreement", "skepticism", and "disagreement", which corresponded to the movement forms we designed with the rotation of three axes. The three different conditions were: when the ear animations appear with no sound; when the ear animations appear simultaneously with voice; and when only voice is heard while no visual image appears. In General discussion, we discussed the possibility of nonverbal communication generated by the combination of our newly designed ear animations and voice. We also discussed the limitations of this research and future challenges.

3. Ear animation design

3.1. Basic guidelines for designing ear animations

We set the following five guidelines for designing the nonverbal communication utilizing ear-movement forms in remote communication where cameras are turned off and participants' faces are not shown.

1. Facial expressions that require individual design of features (eyebrows, eyes, and mouth) are not included.
2. They can convey information in a way that is easy to understand.
3. They can convey emotions in a way that is easy to understand.
4. They give a friendly impression.
5. They are easily noticeable.

As a new rendering that meets all five requirements, we focused on an ear animation representing a rabbit's ears. The animated ears do not need to be limited to those of a rabbit, as they can be the ears of any animal. However, we adopted rabbit ears since their long ears are very expressive. The advantages we can expect from designing nonverbal communication using long ears are as follows:

1. No facial design is required by using only ears for animation.
2. Ears that are placed as if they protrude upward from the top of the head improve visibility.
3. There is a possibility that information, emotion and appearance are expressed with ear movement.
4. The use of rabbit ears gives a friendly impression.

In the overall design, the ears are placed on top of the "main head/body" as if they protrude upward. The shape of head/body in two dimensions is a circle as the shape in three dimensions is a sphere. Any details related to facial features such as eyes, eyebrows, and a mouth are not included in the main head/body, so that it looks like a featureless face. The main body represents a human head and body combined. It functions to display the identity of its corresponding speaker. The identification includes the speaker's name, initial, position (host, participant), and organization. The shape is not limited to a circle or sphere. Because this study is for primary research, a simple sphere was used for the main head/body. Fig. 1 shows a basic image of the ear animations which consists of the main head/body and ears. The ear animations that appeared on the monitor screen as visual stimulus samples did not show any text such as initials that could identify a person. The color and texture of the ear animations can act as an identification information unique to its user. However, since this experiment was not to evaluate individual design but to evaluate the impression of generalized ear movement response, a simple design was used.

As far as we know, this is the first research that uses ear animations with no facial features in a central role for nonverbal communication in remote communication.

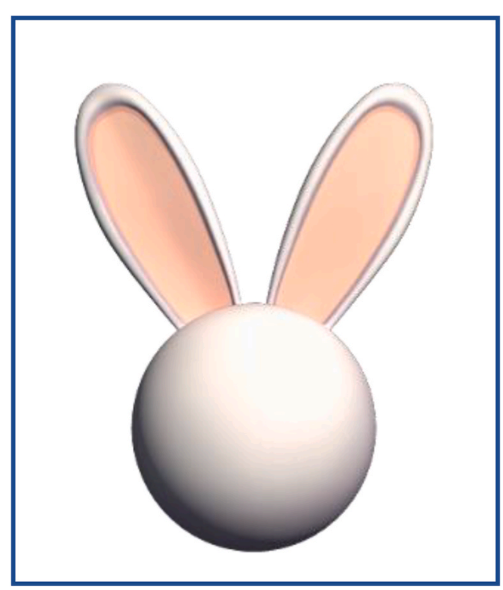


Fig. 1. Basic form of ear animation (primary state prior to start moving).

3.2. Specific design of ear animation

In order to extract specific emotions in remote communication, we referred to two previous studies as guidance for designing the ear animation. The first study is about the Specific Affect Coding System (Gottman, McCoy, Coan, & Collier, 1996; Harrigan et al., 2005). A total of 20 different codes are used as specific affect codes for nonverbal communication between couples in a contracted relationship to encode and categorize the codes into positive, negative and neutral. However, these specific affect codes for nonverbal communication are intended for evaluating interaction between couples in intimate relationships, such as committed couples including married couples. Therefore, not all of the specific affect codes necessarily match those for a professional meeting, which is our target. In order to set suitable codes for our target, we adopted the viewpoint of the other previous study: an output communication/persuasion matrix that reported 12 steps as listener responses for persuasion psychology (McGuire, 1985). In this study, 12 steps were set for a listener's outputs. As our study focused on remote communication meetings, three to seven steps particularly corresponded. Based on these two previous studies, we designed the following six codes as basic targets for nonverbal communication from the viewpoint of listener reaction in business meetings. These are: 1) agreement and 2) applause as a positive code, 3) neutral and 4) confusion as a neutral code, 5) skepticism and 6) disagreement as a negative code.

For designing the ear animation's movement corresponding to the nonverbal communication selected, we referred to a human's head movements and gestures. The meanings of head movements as nonverbal information have been revealed with research on detecting head movements (Maynard, 1987; Buján, 2019). For example, validation is expressed by "nodding", a slow back-and-forth head movement. Critical expression is expressed by "neck swiveling", a side-to-side head tilting movement. Disgust or disagreement is expressed by "shaking", a head turning movement.

Fig. 2 shows the ear animation's basic form and its rotation movement axes: X, Y, and Z. "Agreement" expresses positive nonverbal communication indicating agreement. It is represented by pitching, or X axis rotation which is a head-bending movement. "Skepticism" is designed to express mild negative nonverbal communication. We represented it by rolling, or Y axis rotation which is a head-tilting movement. "Disagreement" expresses a strong negative nonverbal impression indicating rejection. We designed it to correspond by yawing, or Z axis rotation where the head's vertical axis is rotating. "Applause" is designed as the ears clapping. It calls up an image of clapping hands. For "neutral", ears contract vertically to express a neutral state where one cannot make a decision. "Confusion" is expressed by the ears entwining, showing it is hard to understand. Fig. 3 shows the list of movement forms (typical drawing) for the ear animations we designed and the corresponding emotions we intended in their designs. The movement forms describe from their basic (beginning) posture to their posture in the middle, then to the final posture. We adjusted the speed of head movement and hand clapping movement to be the same so that it gives a natural impression. Regarding ear movement and the speed of the ear animations designed, two public licensed psychologists observed the animations in advance and confirmed the animations did not have any particular issue as animations in dialogue interaction. Videos that show 6 different ear movements in the ear animations are presented at: https://github.com/core-dx/mimichara/tree/main/paper/videodata/movement.

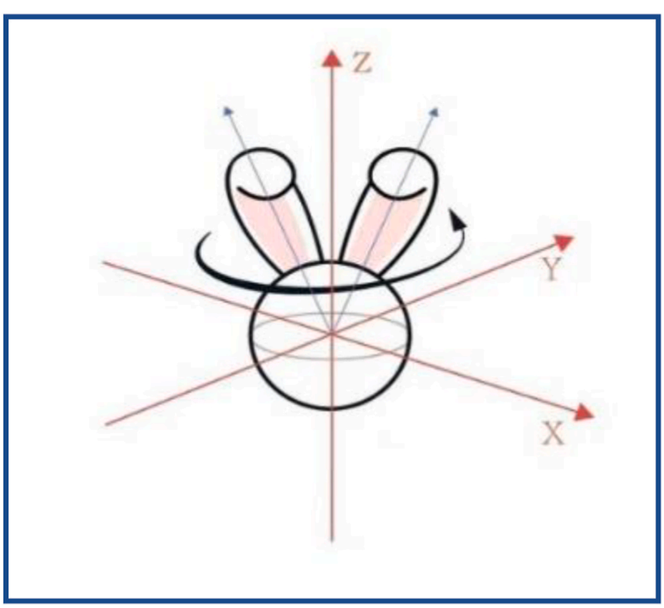


Fig. 2. Relationship among rotation movement axes: X, Y, and Z of the ear animation.

4. Experiment 1: evaluating user impressions of ear-animation-movement forms

4.1. Purpose of Experiment 1

We designed ear movement forms to act as nonverbal communication needed in remote communication by replacing the typical head movements and hand gestures of Japanese people with ear animations. Would these ear animations maintain consistency with our intention of design? The purpose of Experiment 1 was to rate impressions of nonverbal information conveyed by movement of the ear animations. In the experiment, we conducted six levels of one-way experiment planning (repeated) by setting six types of ear animation movement we designed as independent variables, and the impression ratings of the ear animation's movements as dependent variables.

4.2. Hypothesis for Experiment 1

The hypothesis for Experiment 1 is as follows:

H1. Ears-bending-forward movement gives an impression of agreement.

H2. Ears-clapping movement gives an impression of applause.

H3. Ears-contracting movement gives a neutral impression.

H4. Ears-entwining movement gives an impression of comprehension difficulty.

H5. Ears-leaning-to-one-side movement gives an impression of skepticism.

H6. Ears-turning movement gives an impression of disagreement.

4.3. Method

4.3.1. Structure and procedure of Experiment 1

In Experiment 1, we attempted to validate H1 to H6 by seeing whether each ear-animation-movement form corresponded to its intended impression. For the validation, each ear movement was shown to the participants. Then we asked them six questions regarding their impressions from the ear movements. Six-point likert-type scales were used for the questionnaires by setting a 6-point scale from Never: 0 to Yes, Strongly Agree: 5. Participants watched videos of the ear-animation-movement forms, then filled in their answers for the questions in Table 1.

Prior to the impression rating of the ear-animation-movement forms, we obtained information regarding the gender and the age of each participant. Each participant sat in front of a display, which was an 11-inch iPad in portrait mode. The distance to the display was

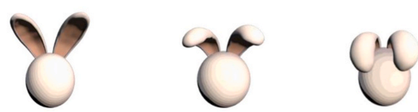
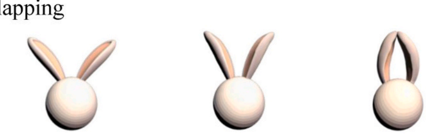
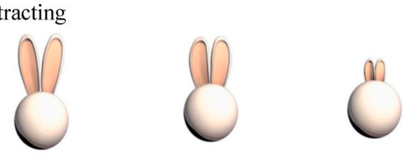
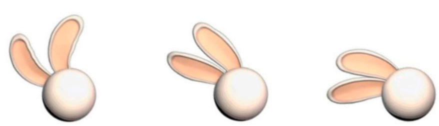
Illustrations for ear movement transition in the ear animation from basic (beginning) posture to posture in the middle, then to final posture	Movement description	Content intended in design
Ears bending forward 	Ears bend forward deeply from central part once	Agreement
Ears clapping 	Both ears show clapping movement, repeating three times	Applause
Ears contracting 	Both ears close to contract once	Neutral
Ears entwining with one another 	Both ears entwine with one another once	Confusion
Ears leaning to one side 	Both ears leaning 90 degrees to one side once	Skepticism
Ears turning 	Both ears turn +/- 65 degrees from side to side once	Disagreement

Fig. 3. The list of ear animation movement forms, movement description and the content intended in design.

Table 1
Experiment 1: Questions for evaluating impressions of ear-animation-movement forms.

1. Did you feel the movement expressing agreement, was like saying "I see"?
2. Did you feel the movement expressing applause, was like saying "Great"?
3. Did you feel the movement expressing comprehension difficulty, was like saying "I don't get it"?
4. Did you feel the movement expressing a neutral attitude, was like saying "Neither"?
5. Did you feel the movement expressing skepticism, was like saying "I'm not sure about that"?
6. Did you feel the movement expressing disagreement, was like saying "I don't think so"?

approximately 50 cm. The ear-animation-movement forms were displayed on the upper side of a white screen. We formatted each image to a size of 50 mm × 60 mm. The video for each ear-animation-movement form was displayed without any accompanying sound. Each movement was shown for 1.5 s. Once completed, there was a 2-s interval, then the same ear-movement form was displayed repeatedly. The questions and the scale were shown underneath the ear-animation-movement forms, so that the participants input their answers by clicking the number on the scale corresponding to their rating. Once all of the answers were filled out, the NEXT button located at the lower part of the screen was activated. By clicking the activated NEXT button, the screen automatically transitioned to the next ear animation. The display order of ear-animation-movement forms was determined by setting a counterbalance and combining ascending and descending methods to avoid an order effect. Once answers for all ear-animation-movement forms were provided and the NEXT button was clicked, a blank frame for free description appeared so that a participant could input a comment freely on the tablet device. We also interviewed the participants to collect their comments. The numerical values of the answers given to the questions were processed with SPSS to compile the data for statistical testing. The experiment consisted of observing six patterns of movement and answering six questions, making a total of 36 items to be judged. It took approximately 10 min for a participant to complete the whole experiment including filling out the answer to the the open answer questions.

In Experiment 1, we used an in-subject experiment. Using G*Power (Faul et al., 2009), we estimated the sample size required for reproducing a large effect size (f = 0.4; 1-β = 0.80) since head movements provide distinct movement differences. Then we considered a study consisting of one group and six ear-movement forms. According to this analysis, at least 11 participants were required.

4.3.2. Participants

The research experiment received full ethical approval from the Psychological Research Ethic Committee at the University of Tsukuba, Japan prior to its commencement. The criteria for participants were as follows. They live and work in Japan and use remote meetings such as videoconferencing. They speak and write Japanese in daily life and at work. They were recruited as opportunity samples without any recompense. No particular exclusion criteria were set for participants. Prior to the evaluation of impressions of the ear animation movements, we obtained the participants' gender and age. The final sample consisted of 24 participants. Their average age was 45.04 (SD = 11.91, the youngest 21, the oldest 62) while 17 of the participants were male and 7 of them were female.

4.4. Results of Experiment 1

In order to examine the impressions of the ear-animation-movement forms, we conducted a one-way analysis of variance (repeated) to analyze the impressions of the following six different ear-animation-movement forms. Figs. 4 and 5 show the results of Experiment 1. In Fig. 4, the letters in bold font with underline indicate the average value of impressions as intended in each movement design, while the result of multiple comparisons is shown by numbers that indicate the significance of each intended design.

1. For the movement of ears bending forward, the following main effect of impression was observed ($F(5,138) = 129.97, p < 0.001$, *partial* $\eta^2 = 0.825$). Bonferroni post-hoc (5%) showed a significant difference was observed in "agreement" as intended in design versus all the other movements.
2. For the movement of ears clapping, the following main effect of impression was observed ($F(5,138) = 121.87, p < 0.001$, *partial* $\eta^2 = 0.819$). Bonferroni post-hoc (5%) showed a significant difference was observed in "applause" as intended in design compared to all the other movements.
3. For the movement of ears contracting, the following main effect of impression was observed ($F(5,138) = 16.029, p < 0.001$, *partial* $\eta^2 = 0.367$). Bonferroni post-hoc (5%) showed that a significant difference was observed in "neutral" as intended in design versus "agreement", "applause" and "disagreement". No significant difference was found versus "confusion" and "skepticism".
4. For the movement of ears entwining, the following main effect of impression was observed ($F(5,138) = 12.536, p < 0.001$, *partial* $\eta^2 = 0.312$). Bonferroni post-hoc (5%) showed that a significant difference was observed in "confusion" as intended in design versus "agreement", "applause", and "disagreement". No significant difference was found versus "neutral" and "confusion".
5. For the movement of ears leaning to one side, the following main effect of impression was observed ($F(5,138) = 25.894, p < 0.001$, *partial* $\eta^2 = 0.486$). Bonferroni post-hoc (5%) showed that a

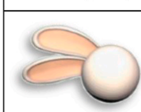
	User Impression Rating					
Ear movement form	Agreement 1	Applause 2	Neutral 3	Confusion 4	Skepticism 5	Disagreement 6
 1.58(2.00)	**<u>4.79</u>** (0.42) 1>2,3,4,5,6	3.33 (1.61) 2>3,4,5,6 2<1	0.46 (0.78) 3<1,2	0.33 (0.64) 4<1.2	0.25 (0.53) 5<1,2	0.29 (0.55) 6<1,2
 1.61(2.09)	3.62 (1.74) 1>3,4,5,6 1<2	**<u>4.83</u>** (0.38) 2>1,3,4,5,6	0.37 (0.65) 3<1,2	0.25 (0.53) 4<1,2	0.29 (0.55) 5<1,2	0.29 (0.86) 6<1,2
 1.88(1.80)	0.75 (0.94) 1<3,4,5	0.63 (1.06) 2<3,4,5	**<u>2.54</u>** (1.61) 3>1,2,6	2.62 (1.71) 4>1,2,6	3.54 (1.53) 5>1,2,6	1.21 (1.67) 6<3,4,5
 1.87(1.89)	0.54 (1.06) 1<3,4,5	0.50 (0.72) 2<3,4,5,6	2.37 (1.71) 3>1,2	**<u>3.29</u>** (1.78) 4>1,2,6	2.63 (2.00) 5>1,2	1.88 (1.78) 6>2 6<4
 2.33(2.05)	0.71 (1.33) 1<3,4,5,6	0.63 (1.31) 2<3,4,5,6	2.17 (1.79) 3>1,2 3<4,5	3.92 (1.47) 4>1,2,3,6	**<u>4.33</u>** (1.31) 5>1,2,3,6	2.25 (1.70) 6>1,2 6<4,5
 1.86(1.84)	0.71 (1.08) 1<4,5,6	0.46 (0.93) 2<3,4,5,6	1.88 (1.68) 3>2 3<4	3.38 (1.61) 4>1,2,3	2.13 (1.80) 5>1,2	**<u>2.63</u>** (1.95) 6>1,2

Fig. 4. Results of the experiment in ear-movement forms and impression evaluation (average, (SD), the result of multiple comparisons is indicated as a number below each impression, the bold-underlined value in each animation corresponds to our intended design).

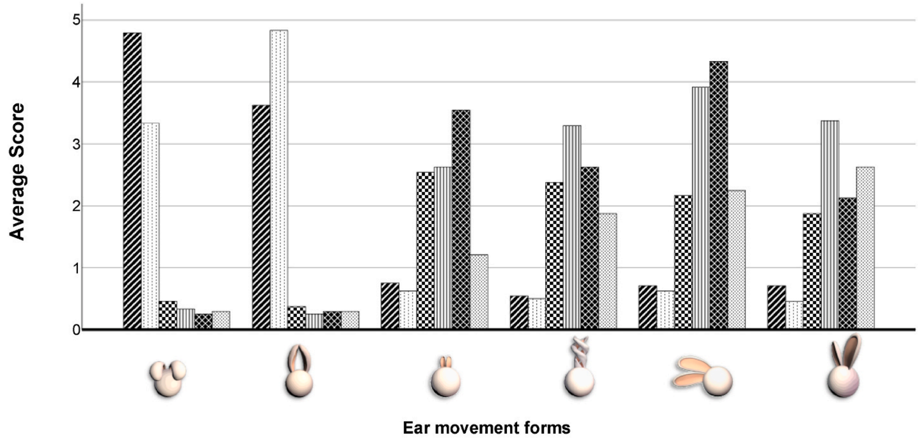


Fig. 5. Evaluation of impressions given by the ear-animation-movement forms.

significant difference was observed in "skepticism" as intended in design versus "agreement", "applause", "neutral" and "disagreement". No significant difference was found versus "confusion".

6. For the movement of ears turning, the following main effect of impression was observed ($F(5,138) = 12.401$, $p < 0.001$, *partial* $\eta^2 = 0.310$). Bonferroni post-hoc (5%) showed that a significant difference was observed in "disagreement" as intended in design in comparison to "agreement", "applause", and "neutral". No significant difference was found versus "neutral", "confusion" and "skepticism".

4.5. Discussion of Experiment 1

For ears bending forward and ears clapping, "agreement" and "applause" supported hypotheses H1 and H2. We were able to confirm user impressions on the "agreement" and "applause" ear animations as intended in design, resulting from those movements being natural enough to be easily comprehensible according to open answers and interviews. For ears contracting, ears entwinning, ears leaning to one side and ears turning, the respective impressions they received were not the ones we intended in design. Therefore, hypotheses H3, H4, H5 and H6 were not supported. In particular, the "disagreement" animation did not reflect our design intention. Since the turning angles of the head movement were $\pm$ 65° with a fast-turning speed, many participants commented in their open answers that it gave the impression that it was looking around restlessly or it was looking for something. It is likely that the set values for the head turning angles and speed were too high for Japanese people to perceive the movements. Also, the "disagreement" ear animation movement where the vertical axis of the head rotates, gave the participants the impression of "confusion" or "skepticism". Although the animation was able to give the impression of a negative movement, intensity (angles) and speed of the movement are important for conveying subtle intention. This also indicates the limitation of impressions given to the participants. The ear animation intended for "neutral" ended up giving an impression of "confusion". For this reason, participants commented in their open answers or interviews that it was difficult to understand the movement because it was the first time for them to see such a movement. Since this movement was not a simple movement like turning a head, it seemed to be difficult to judge. It is likely that the ears-entwining movement for "confusion" was affected by its difference to natural movements as the participants formed their impression of it. While the movements had difficulty in conveying subtle impressions, they showed significant effects on classifying the movements as negative or positive. This indicates that movements that do not exist in nature can give a positive or negative impression. This finding can be interpreted to mean that we will have a certain degree of freedom for conveying subtle impressions if an ear animation is displayed with words, or if we define the meaning of each movement in advance. It suggests the possibility that both enhancing and softening impressions can be done.

Furthermore, there were comments in the open answers describing the ear movements as cute and that they looked friendly. The participants had positive impressions in regard to how the ear animations were presented.

5. Experiment 2: evaluation of user impressions in condition where simple voice and ear animation presented simultaneously

5.1. Purpose of Experiment 2

In Experiment 1, we compared the ratings of user impressions given by the six types of ear animation movement forms. In Experiment 2, we added impression ratings based on conveying method as an independent variable. There are three conditions in the conveying methods: when ear animations were presented with no sound, when ear animations were presented simultaneously with simple voice, when only voice was played. Under these three conditions, we compared psychological impressions. In addition, as independent variables, we set three kinds of content to be conveyed: "agreement", "skepticism", and "disagreement" corresponding to three movement forms: ears bending forward, ears leaning to one side, and ears turning, all of which followed rotation movement on axes: X, Y, and Z of the ear animation shown in Fig. 2. This was a Two-way Repeated-Measures in which conveying method and content conveyed were set as independent variables. Dependent variables were conveying information and conveying emotions. This experiment was to rate impression evaluation from these two points of view. The reason why we added movement of contents conveyed as an independent variable was to review any influence from voice assigned to each different ear animation. In Experiment 2, we chose "agreement" which gave strong and clear impressions and supported our hypothesis in Experiment 1, "skepticism" which gave a relatively strong impression in the experiment's impression rating, and "disagreement" which gave a vague impression in the rating. Having these three types of movement forms as the independent variables, we observed how the effect of adding voice to a movement can affect the intensity of impression over the silent movement alone. To explore the possibility of the ear animations being used in a remote communication application, we also conducted questionnaires asking if the participants felt the ear animation would be promising application.

5.2. Hypothesis for Experiment 2

If voice and the ear animation are played simultaneously, there will be integration of auditory and visual information, and we can expect this integration to make it easier to convey information and emotion. For content conveyed, the movement forms of the ear animation are considered to be factors for conveying information and emotion. In Experiment 1, "agreement" (ears bending forward) scored significantly higher impression rates than "skepticism" (ears leaning to one side) and "disagreement" (ears turning). Therefore, we can expect that interaction between conveying method and content conveyed will be generated. We can also expect that the combination of the ear animation with voice (voice & movement) and content conveying "agreement" will score higher than other combinations. Since ear animations with voice are expected to score higher than the voice-only communication, we expect the ear animations with voice will be utilized in remote communication where nonverbal information can be added. Hypotheses for this research are as follows:

H7. In conveying information, there is interaction between conveying method and content conveyed, and the combination of an ear animation with voice (voice & movement) as a conveying method and "agreement" as a content conveyed scores higher than the other combinations.

H8. In conveying emotion, there is interaction between conveying method and content conveyed, and the combination of an ear animation with voice (voice & movement) as a conveying method and "agreement" as a content conveyed scores higher than the other combinations.

H9. An ear animation with voice is promising in remote communication.

5.3. Method

5.3.1. Structure and procedure of Experiment 2

The structure and procedure for Experiment 2 are basically the same as for Experiment 1. The same ear animations for "agreement", "skepticism" and "disagreement" as in Experiment 1 were used. We produced voice sounds using voice synthesis software to correspond to each animation. For expressing agreement, the voiced word "I see" was uttered. For expressing skepticism, the voiced word "I'm not sure about that" was uttered. For expressing disagreement, the voiced word "I don't think so" was uttered. As all these voiced words were actually spoken in Japanese, we translated them to English in this paper for convenience. Videos that show the ear movements with voice (spoken in Japanese) are presented at https://github.com/core-dx/mimichara/tree/main/paper/videodata/movementvoice. In Experiment 2, we compared evaluations of dialogue responses for three different conveying methods: ear animations with voice (voice & movement), silent ear animations (movement-only), and only audio (voice-only). A male voice was used for voice synthesis while the length of each utterance was set to 1.5 s or shorter, so as to fit the length of time for each corresponding ear animation movement.

As in Experiment 1, each participant sat in front of a display, which was an 11-inch iPad placed in portrait mode. The distance to the display was approximately 50 cm. The ear-animation-movement forms were displayed on the upper side of a white screen. We formatted each image to a size of 50 mm × 60 mm. The method to display the animations, present questions, the criteria, the open form question, and interview were all the same as the ones for Experiment 1. For the voice-only method, there was nothing displayed on the 50 mm × 60 mm display. Voice was output from the iPad's speaker. Table 2 lists question items for Experiment 2. The experiment consisted of observing the three levels of content conveyed: "agreement", "skepticism" and "disagreement" and three conveying methods: voice-only, silent ear-animation (movement-only) and ear animation with voice (voice & movement); and answering three questions for each, making up 27 items for evaluation. It took approximately 10 min for a participant to complete the whole experiment including filling out the answer to the open form question.

In Experiment 2, we used an in-subject experiment. Using G*Power (Faul et al., 2009), we estimated the sample size required for reproducing a large effect size (f = 0.4; 1-β = 0.80), since head movements provide distinct movement differences. Then we considered a study consisting of one group, and nine measurements based on three levels multiplied by three levels. According to this analysis, at least 9 participants were required.

5.3.2. Participants

The research experiment received full ethical approval from the Psychological Research Ethic Committee at the University of Tsukuba, Japan prior to its commencement. Participants in Experiment 1 participated in Experiment 2 as well. The final sample consisted of 24 participants. Their average age was 45.04 (SD = 11.91, the youngest 21, the oldest 62) while 17 of the participants were male and 7 of them were female.

5.3.3. Results of Experiment 2

To carry out statistical analysis of the ear animation with voice (voice & movement), we conducted Two-way ANOVA (repeated) setting conveying method and content conveyed as independent variables to analyze conveying information and conveying emotion respectively.

5.3.3.1. The result of Two-way ANOVA (repeated), conveying information as a dependent variable. The main effect of the contents conveyed ($F(2,46) = 5.60$, $p < 0.01$, $\eta^2 = 0.24$) and the main effect of the conveying method ($F(2,46) = 9.42$, $p < 0.01$, $\eta^2 = 0.41$) were significant. Also, the interaction effect between the content conveyed and the conveying method ($F(4,92) = 6.80$, $p < 0.01$, $\eta^2 = 0.30$) was significant. We conducted simple main effect analysis on these interaction effects.

1. The simple main effect of the content conveyed for the ear animation with voice (voice & movement) showed marginal significance ($F(2,46) = 2.74$, $p < 0.1$, $\eta^2 = 0.12$). When Bonferroni multiple comparison was conducted (5%), no significance was observed among the content conveyed for the ear animations with voice (voice & movement).
2. The simple main effect of the content conveyed for the silent ear animation (movement-only) showed significance ($F(2,46) = 6.94$, $p < 0.01$, $\eta^2 = 0.30$). When Bonferroni multiple comparison was conducted (5%), for the content conveyed with the silent ear animation (movement-only), "agreement" scored significantly higher compared to "disagreement".
3. The simple main effect of the content conveyed with voice-only showed significance ($F(2,46) = 7.36$, $p < 0.01$, $\eta^2 = 0.30$). When Bonferroni multiple comparison was conducted (5%), for the content conveyed with voice-only, "agreement" scored significantly higher

Table 2
Experiment 2: Questions for evaluating the communicative characteristics of dialogue responses using ear animations.

1. Did you feel this response conveyed the information from your partner?
2. Did you feel this response conveyed your partner's emotion?
3. Did you feel this way of conveying a response could be used effectively in remote meetings?

compared to "skepticism", while "disagreement" scored significantly higher compared to "skepticism".

4. The simple main effect of the conveying method for "agreement" showed significance ($F(2,46) = 6.07$, $p < 0.01$, $\eta^2 = 0.26$). When Bonferroni multiple comparison was conducted (5%), for the conveying method of "agreement", the ear animation with voice (voice & movement) scored significantly higher compared to the silent ear animation (movement-only) and voice-only.
5. The simple main effect of the conveying method for "skepticism" showed significance ($F(2,46) = 7.47$, $p < 0.01$, $\eta^2 = 0.32$). When Bonferroni multiple comparison was conducted (5%), for the conveying method for "skepticism", the ear animation with voice (voice & movement) scored significantly higher compared to the silent ear animation (movement-only) and voice-only.
6. The simple main effect of the conveying method for "disagreement" showed significance ($F(2,46) = 10.30$ $p < 0.01$, $\eta^2 = 0.44$). When Bonferroni multiple comparison was conducted (5%), for the conveying method for "disagreement", the ear animation with voice (voice & movement) scored significantly higher compared to the silent ear animation (movement-only), while voice-only scores significantly higher compared to the silent ear animation (movement-only).

5.3.3.2. The result of Two-way ANOVA (repeated), conveying emotion as a dependent variable. The main effect of the content conveyed ($F(2,46) = 3.42$, $p < 0.05$, $\eta^2 = 0.15$) and the main effect of the conveying method ($F(2,46) = 12.54$, $p < 0.01$, $\eta^2 = 0.54$) were significant. Also, the interaction effect between the content conveyed and the conveying method ($F(4,92) = 6.29$, $p < 0.01$, $\eta^2 = 0.27$) was significant. We conducted simple main effect analysis on this interaction effect.

1. The simple main effect of the content conveyed with the ear animation with voice (voice & movement) showed significant difference ($F(2,46) = 5.17$, $p < 0.05$, $\eta^2 = 0.22$). When Bonferroni multiple comparison was conducted (5%), for the content conveyed with the ear animation with voice (voice & movement), "agreement" scored significantly higher compared to "disagreement".
2. The simple main effect of the content conveyed with the silent ear animation (movement-only) showed significance ($F(2,46) = 9.73$, $p < 0.01$, $\eta^2 = 0.42$). When Bonferroni multiple comparison was conducted (5%), for the content conveyed with the silent ear animation (movement-only), "agreement" scored significantly higher compared to "disagreement", while "skepticism" scored higher compared to "disagreement".
3. The simple main effect of the content conveyed with voice only did not show significance.
4. The simple main effect of the conveying method for "agreement" showed significance ($F(2,46) = 12.26$, $p < 0.01$, $\eta^2 = 0.57$). When Bonferroni multiple comparison was conducted (5%), for the conveying method for "agreement", the ear animation with voice (voice & movement) scored significantly higher compared to the silent ear animation (movement-only) and voice-only.
5. The simple main effect of the conveying method for "skepticism" showed significance ($F(2,46) = 8.02$, $p < 0.01$, $\eta^2 = 0.35$). When Bonferroni multiple comparison was conducted (5%), for the conveying method for "skepticism", the ear animation with voice (voice & movement) scored significantly higher compared to the silent ear animation (movement-only) and voice-only.
6. The simple main effect of the conveying method for "disagreement" showed significance ($F(2,46) = 8.76$ $p < 0.01$, $\eta^2 = 0.38$). When Bonferroni multiple comparison was conducted (5%), for the conveying method for "disagreement", the ear animation with voice (voice & movement) scored significantly higher compared to the silent ear animation (movement-only), while voice only scores significantly higher compared to the silent ear animation (movement-only).

Table 3 shows the results of the evaluation of user impressions in Experiment 2.

5.3.3.3. Results of user expectation on ear animations (voice & movement) to be used in remote communication. Regarding user expectations of the ear animations to be used in remote communication, a main effect was observed ($F(2,207) = 17.057$ $p < 0.001$, *partial* $\eta^2 = 0.141$) for the conveying method. In the results of the Bonferroni post-hoc, the conveying method using ear-animations with voice was observed as more significant than both the method using voice-only and the one using silent ear-animation. For content conveyed, a main effect was observed ($F(2,207) = 4.957$ $p < 0.001$, *partial* $\eta^2 = 0.046$). In the results of Bonferroni post-hoc (5%), for conveying information, a significant difference was observed in “agreement” versus “disagreement”. No interaction was observed between content conveyed and conveying method ($F(4,207) = 1.411$ $p < 0.231$, *partial* $\eta^2 = 0.027$). Table 4 shows the results of evaluation of user expectation on remote communication in Experiment 2.

5.3.4. Discussion of Experiment 2

Interaction between a conveying method and content conveyed was acknowledged. The combination of the ear animation with voice (voice & movement) and “agreement” was significantly superior to that of silent ear animation (movement-only) and voice-only. This supported hypotheses H7 and H8. It is believed that a synergistic effect of the ear animation’s movement and voice appeared in the conveying method. Many comments in the answer to the open question mentioned that it got much easier to understand what the ear animation wanted to convey when voice was played simultaneously. Interestingly, in Experiment 1, the ears-leaning-to-one-side intended to mean "skepticism" and the ears-turning intended to mean "disagreement" were perceived as "confusion". This indicates even though impressions of these movements are vague when only their movements are shown, communicative characteristics of their movements become clear by adding voice to them.

For content conveyed, "agreement" had significant difference versus "disagreement" and "skepticism" in conveying information and emotion. As revealed in Experiment 1, it is conceivably affected by the fact that the impression of the ear movement was very strong for "agreement". This proves that the ear animation with voice showed significant differences in conveying information in terms of content conveyed as well as impression scoring rate in terms of conveying emotion. It is important to optimize the movement forms of the ear animations with voice.

Regarding the interaction between content conveyed and conveying method, Figs. 6–8 show the relationships between their total scores. Figs. 6 and 7 respectively show the total scores for each condition of conveying method and content-conveyed in regard to conveying information and conveying emotion. In conveying information, the ear animation with voice (voice & movement) was more significant than the movement-only method in regard to "agreement" and "disagreement",

Table 3
Evaluation results of user impressions for content conveyed and conveying method by ear animations (Average value (SD)).

Content conveyed	Conveying method	Information	Emotion
Agreement	Voice-only	3.88(1.33)	2.58(1.82)
	Movement-only	3.83(1.32)	3.75(1.59)
	Voice & Movement	4.67(0.55)	4.58(0.76)
Skepticism	Voice-only	2.92(1.63)	2.71(1.74)
	Movement-only	3.33(1.40)	3.08(1.47)
	Voice & Movement	4.29(1.14)	4.13(1.17)
Disagreement	Voice-only	4.00(1.15)	3.13(1.81)
	Movement-only	2.50(1.63)	2.00(1.55)
	Voice & Movement	4.13(1.27)	3.79(1.32)

Table 4
Results of user expectation on ear animations (voice & movement) to be used in remote communication.

	Conveying method		
Content conveyed	Voice-only	Movement-only	Voice & movement
Agreement	3.25(1.57)	3.63(1.61)	4.42(1.10)
Skepticism	2.21(1.64)	3.17(1.69)	4.17(1.34)
Disagreement	2.71(1.99)	2.29(1.68)	3.92(1.28)

while it was more significant than the voice-only method in regard to "skepticism". In conveying emotion, the ear animation with voice (voice & movement) was more significant than the voice-only method in regard to "agreement" and "skepticism", while it was more significant than the movement-only in regard to "disagreement". When considered in a comprehensive manner, voice is a main contributor for conveying information while movement is a main contributor for conveying emotion. Also, integration of voice and movement seems effective for conveying information and emotion. This indicates the effect on the communicative characteristics of voice and movement are generated without any facial expression. The significant difference from the voice-only method in conveying "skepticism" seems to be caused by the vagueness of the spoken information "I'm not sure about that" in Japanese, since this spoken information is vague compared to the ones for "agreement" and "disagreement". In this experiment, "disagreement" showed a difference from "agreement" and "skepticism" in conveying emotions. As revealed in Experiment 1, it is assumed that the turning movement of the ears, which is the intended design of "disagreement", was weak in conveying impressions. In short, because the impression of the "disagreement" ear-movement form was vague, it is likely that voice and movement became significant communicative characteristics in conveying both information and emotion. In contrast, no significant

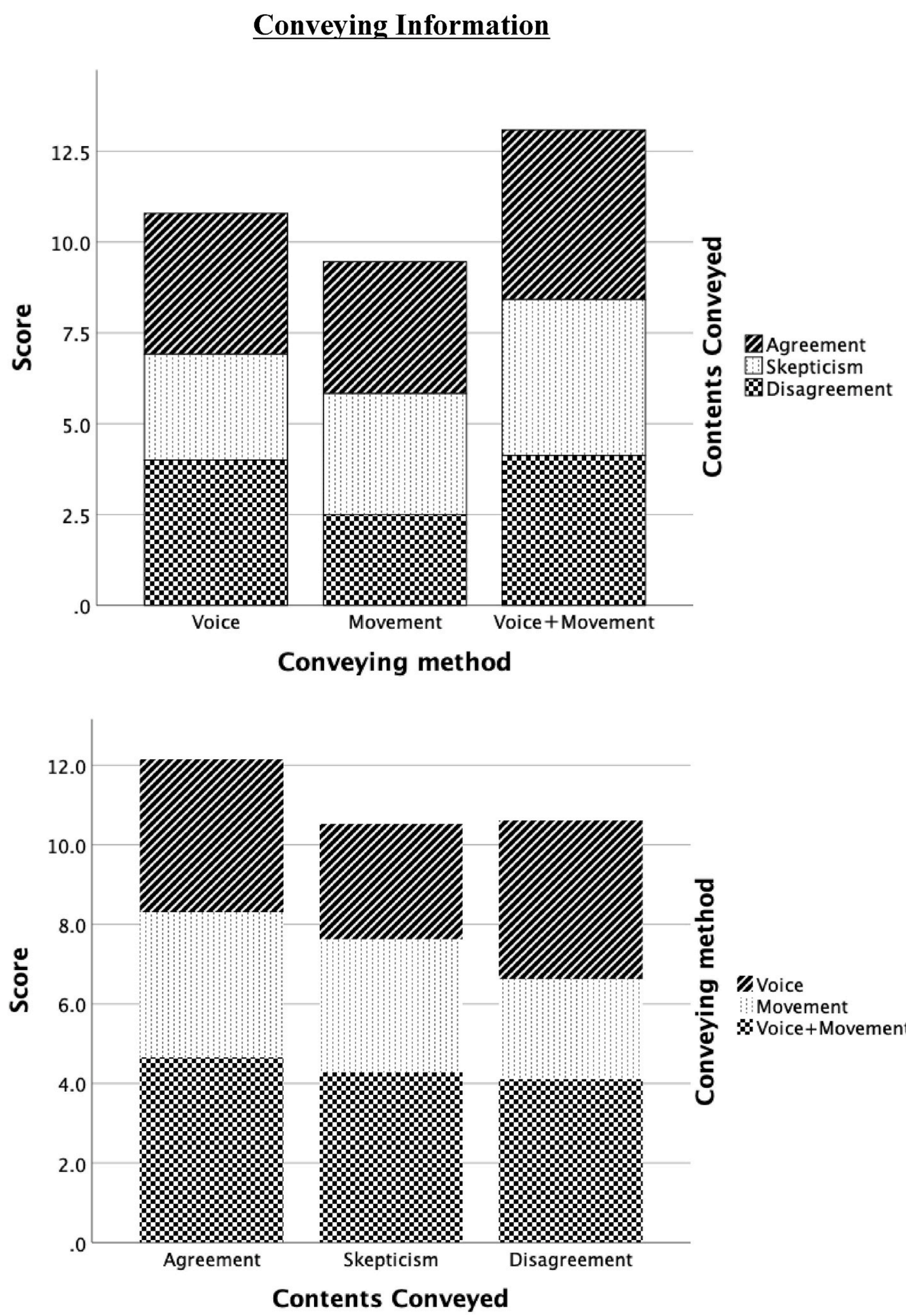


Fig. 6. Evaluation results of content conveyed and conveying method in conveying information.

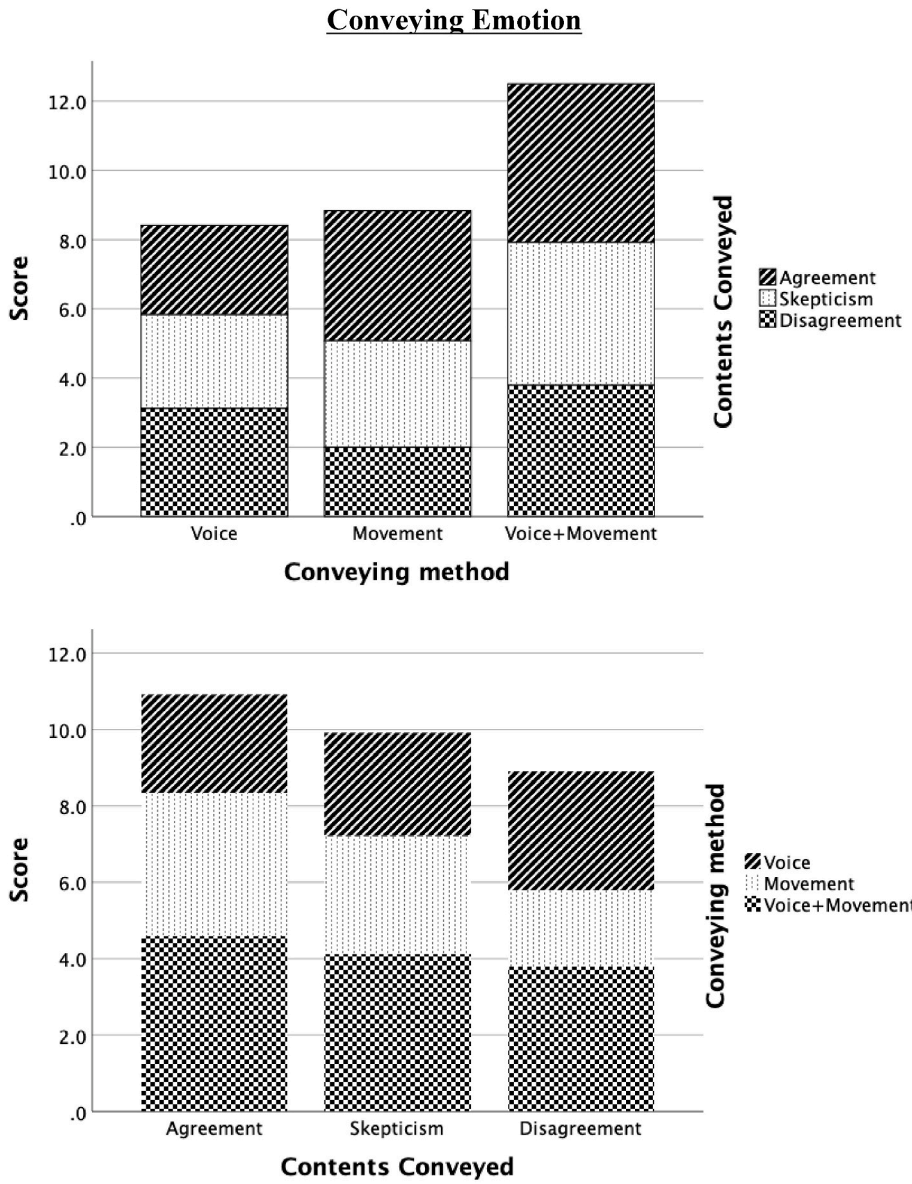


Fig. 7. Evaluation results of content conveyed and conveying method in conveying emotion.

difference was seen between the movement-only and voice-only conveyance methods for "disagreement". This will be improved by reviewing the linguistic expression of "skepticism" and the ear-movement form of "disagreement".

Regarding the high expectancy for ear animations with voice to be used in remote communication, participants felt the ear animations with voice are more effective as a conveying method in all "agreement", "skepticism" and "disagreement" conveyances than the silent ear animations and voice-only. This result supported H9, indicating that we will need to evaluate the ear animations by testing them in remote communication in the future.

6. General discussion

6.1. Discussion

In this study, we conducted primary research on user impressions of featureless-face-ear animations as nonverbal communication that can be used in remote communication where web cameras are turned off. Regarding the design of the ear-animation-movement forms, statistical tests revealed movements close to natural human movements such as "agreement" and "applause" are valid in conveying desired impressions. On the other hand, the "disagreement" movement of ears turning from side to side gave the impression of looking around restlessly. As a result of exaggerating human movements, rather vague impressions were formed. In designing ear-movement forms that could affect content conveyed, it is likely that the intensity of movements and speed adjustment affect the intensity of an impression. Therefore, to design an ear animation, it is necessary to consider the content it wants to convey, movement forms corresponding to its intention, and speed optimization of movement. On the whole, Experiment 2 showed there was interaction between the conveying method and content conveyed in conveying information and conveying emotion. The ear animation and the ear-bending movement of "agreement" scored high in the impression rating. It indicates that it is important to design movement forms for the ear animations relevantly and integrate them with voice. In addition to conventional voice interactions, interactions using video, and interactions using avatars with facial expressions, this study indicated the possibility of a new interaction style using ear animations with no facial information.

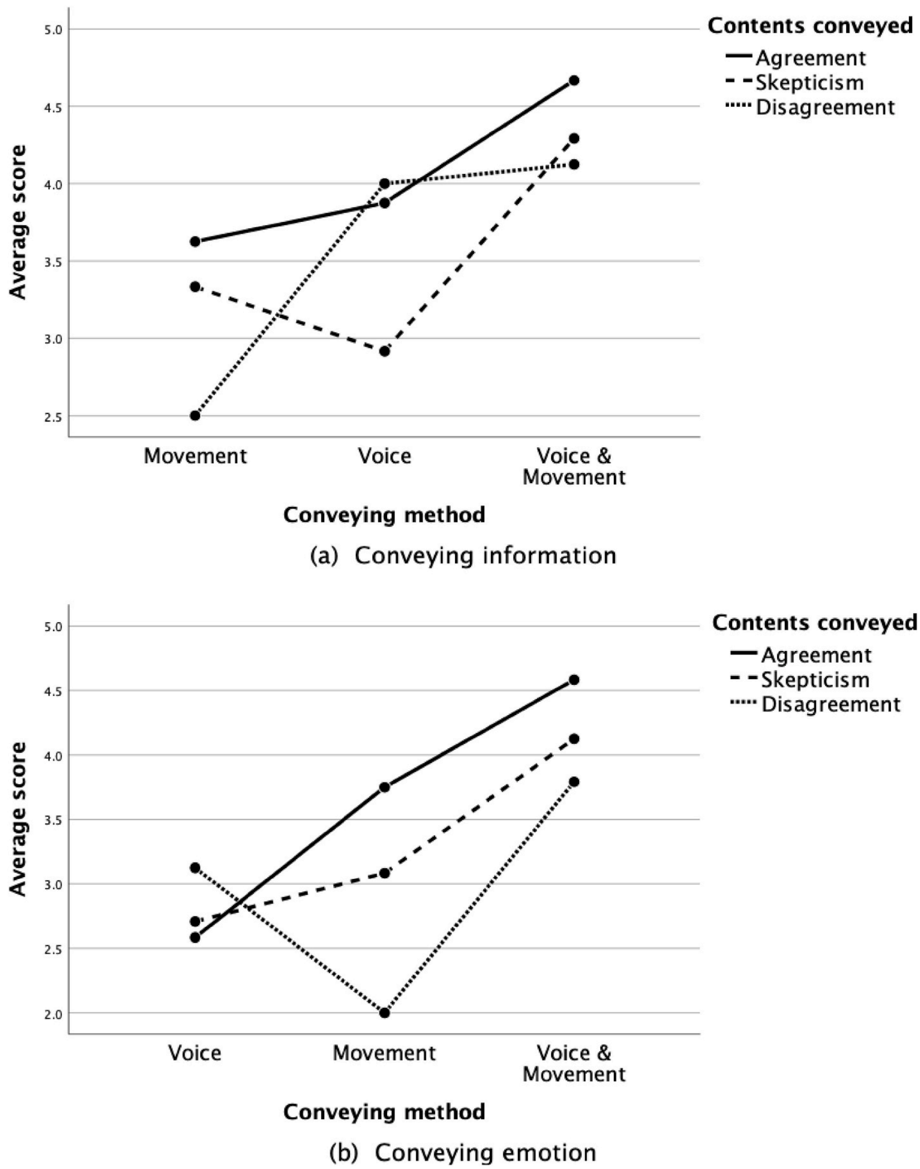


Fig. 8. Average scores of contents conveyed and conveying methods.

In a sense, to use ear animations with no facial expression and no eye gaze is to use nonverbal communication that lacks the most important part of human nonverbal communication. By integrating voice and movement, this communication style is thought to provoke people to use their imagination, making up for the lack of facial expression. That is to say, conveying rich emotions is made possible by the combination of voice and movements even if every detail is not visually expressed as facial information. Minimal design has been proposed as a way to emphasize the main points of communication by eliminating unnecessary design elements (Matsumoto et al., 2006). Moreover, it has been reported that a no-facial-expression robot with voice, which introduced the minimum design method for doll therapy, is positively accepted by elderly adults with dementia (Sumioka et al., 2021). Although this research does not deal with a physical robot, this is considered as a new method of communication that stimulates human imagination with a reduced amount of information by using the movements of the ear animations with no facial expressions.

As previously mentioned, it is difficult to meet and acknowledge another person's gaze over a monitor. This has been a particular issue in real-time video interaction. By utilizing the ear animations, we can expect the issue to be solved because eye gaze itself is completely omitted- in the animation. As long as human imagination can fill in the absence of the information, this might be better than any unnatural eye gaze that occurs on the screen. Contrary to audio-only interaction, the voice and ear animations were positively received for use in remote meetings. It is indirectly supported that the voice and ear animations are accepted even if there is no facial information or eye contact information in them.

Focusing on the ear animations' visual stimuli, we can compare both differences and similarities between the ear animations and emojis. It has been over 20 years since emojis were first introduced. In recent years, the use of emojis has been expanding remarkably as studies on emojis are being conducted actively (Bai et al., 2019). In general, emoticons are used for expressing different emotions (Derks et al., 2008). Many studies have revealed emoticons and emojis enhance enjoyable communication and stimulate arousal levels (Walther & D'Addario, 2001). As relatively simple visual stimuli, the ear animations have similarities to emojis. It was reported that the visual stimuli of emojis are processed faster than a word (Kaye et al., 2021). The ear animations are visual stimuli, but they lack a wide variety of expressions emojis have. To cope with this disadvantage, we are interested in establishing hybrid communication consisting of the ear animations and emojis. This is one of our research themes that we will discuss in the future.

6.2. Limitations of the ear animations in this study

In this study, we designed ear animations representing rabbit ears and evaluated user impressions of nonverbal information such as information and emotion brought by specific movement of the ear animations. While evaluating user impression is affected by various factors such as ear animation design, movement form, and movement speed, this study's evaluation is limited to specific ear animations, not covering evaluations for each factor. In regard to participants, the influence of age, gender, personality, ethnicity, and so on can be considered, but our evaluation was limited to certain groups of people. Although the combined effects of voice and ear animations were suggested in this study, it was limited to combinations of specific words and specific ear animations. Furthermore, we need to consider how to present the ear animations on a screen, such as displaying a single ear animation or multiple ear animations at one time, as well as the observation viewpoint of 3D ear animations. Also, we should not ignore how the back ground of ear animations would affect user impressions. The results of this study are limited to basic and specific examples of ear animations.

6.3. Future challenges for the ear animation

First, we proposed and designed the ear animations, and conducted a primary evaluation of user impressions. For practical use, we need to evaluate the characteristics of dialogues in remote communication using the ear animations.

Second, in remote communication research using the ear animations, how to display a self-image is a future topic of discussion. The ear animations give a sense of existence as avatars in cyberspace. Therefore, it is important to investigate the way ear animations are displayed in cyberspace to form the sense of a communication partner or team unity in remote communication.

Thirdly, although a basic potential for ear animations has been indicated, further research on movement speed, movement intensity and a variety of movement forms is expected in order to improve the communicative characteristics of the ear animations.

7. Conclusion

Through proposing and designing the ear animations, and assessing user impressions of the movement forms and communicative characteristics of dialogue responses, this study suggests that the ear animations that do not contain facial information changing dynamically can be used in remote communication with a web camera turned off to improve overall communication. In particular, "agreement" and "applause" are consistent with our design intentions in evaluating ear-animation-movement forms. Such natural movements easily convey impression and dialogue responses. On the other hand, movements that are unlike natural movements such as "disagreement", "confusion" and "neutral" in this study's designed movement forms easily caused unintended results in conveying emotions. However, the ear movement for "agreement", the ear movements for "skepticism" and the ear movement for "disagreement" were significantly improved by adding voice in both conveying information and emotion compared to voice-only and movement-only. It is likely that the ear animations act as a mechanism with voice producing a synergistic effect of auditory and visual perceptions.

To improve conveying impressions as well as the communicative characteristics of dialogue responses, and to adjust subtle expressions, we comprehensively evaluated the results of the statistical testing and answers to the open questions. Through these evaluations, we will consider further design of the ear-movement forms and adjust their intensity (degree of dynamic movement), speed, and so on.

The ear animations in this study are not equipped with facial expressions changing dynamically. It is considered that the lack of facial information is complemented by human imagination. Because the ear animation design does not require subtle facial expression to be customized, it allows relatively fewer sensors to be used. Therefore, it is easy to implement the ear animations in remote communication. As a result of this study, the ear animations are expected to be applied in remote communication. Going forward, it is desirable to evaluate communication using the ear animations.

Declaration of competing interest

The authors have no conflicts of interest directly relevant to the content of this article. Regarding conflict of interest, the authors receive full approval from the Psychological Research Ethic Committee at the University of Tsukuba, Japan.

Acknowledgements

The authors would like to thank Ms. Tomomi Kosaka and Mr. Michael Gray for assistance with English translation of this paper.

References

Archer, D. M., & Akert, R. M. (1997). Words and everything else: Verbal and nonverbal cues in social interpretation. *Journal of Personality and Social Psychology, 35*, 443–449.

Bai, Q., Dan, Q., Mu, Z., & Yang, M. (2019). A systematic review of emoji: Current research and future perspectives. *Frontiers in Psychology, 10*, 2221, 2019.

Bailenson, J. (2018). *Experience on demand: What virtual reality is, how it works, and what it can do*. New York, USA: W.W. Norton.

Bohannon, L. S., Herbert, A. M., Pelz, J. B., & Rantanen, E. M. (2013). Eye contact and video- mediated communication: A review. *Displays, 34*(2), 177–185. https://doi.org/10.1016/j.displa.2012.10.009CrossRefGoogle Scholar

Buján, M. M. (2019). The function of face gestures and head movements in spontaneous humorous communication. *The European Journal of Humor Research, 7*(2), 1–29. https://doi.org/10.7592/EJHR2019.7.2.bujan

Campbell, A. G., Holz, T., Cosgrove, J., Harlick, M., & O'sullivan, T. (2019). Uses of virtual reality for communication in financial services: A case study on comparing different telepresence interfaces: Virtual reality compared to video conferencing. *Lecture Notes in Networks and Systems Advances in Information and Communication, 2* (8), 463–481, 2019.

Castelli, F., & Sarvary, M. (2021). Why students do not turn on their video cameras during online classes and an equitable and inclusive plan to encourage them to do so. *Academic Practice in Ecology and Evolution, 11*(Issue 8). https://doi.org/10.1002/ece3.7123

Derks, D., Bos, A. E. R., & Grumbkow, J. V. (2008). Emoticons in computer-mediated communication: Social motives and social context. *CyberPsychology and Behavior, 11*, 99–101, 2008.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149–1160.

Forsberg, N. B., & Kirchner, K. (2021). The perception of avatars in virtual reality during professional meetings. In *HCI international 2021 – posters front. Robot. AI, 17 june 2021*. https://doi.org/10.3389/frobt.2021.633378

Gottman, J. M., McCoy, K., Coan, J., & Collier, H. (1996). *The specific affect coding system (SPAFF) for observing emotional communication in marital and family interaction*. Mahwah, NJ: Erlbaum.

Haase, R. F., & Tepper, D. T., Jr. (1972). Nonverbal components of empathic communication. *Journal of Counseling Psychology, 19*, 417–426.

Harrigan, J., Rosenthal, R., & Scherer, K. R. (Eds.). (2005). *The new handbook of methods in nonverbal behavior research*. Oxford: Oxford University Press.

Junuzovic, S., Inkpen, K., Tang, J., Sedlins, M., & Fisher, K. (2012). To see or not to see: A study comparing four-way avatar, video, and audio conferencing for work. In *Proceedings of the 17th ACM international conference on Supporting group work GROUP '12* (pp. 31–34).

Kaye, L. K., Rodriguez-Cuadrado, S., Malone, S. A., Wall, H. J., Gaunt, E., Mulvey, A. L., & Graham, C. (2021). How emotional are emoji?: Exploring the effect of emotional valence on the processing of emoji stimuli. *Computers in Human Behavior, 116*, Article 106648.

Matsumoto, N., Fujii, H., & Okada, M. (2006). Minimal design for human-agent Communication. *Artificial Life and Robotics, 10*(1), 49–54.
Maynard, S. K. (1987). Interactional functions of a nonverbal sign: Head movement in Japanese dyadic casual conversation. *Journal of Pragmatics, 11*(5), 589–606.
McGuire, W. J. (1985). Attitudes and attitude change. In G. Lindzey, & E. Aronson (Eds.), *Handbook of social psychology* (3rd ed., Vol. 2). New York Random House.
Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley. *Robotics & Automation Magazine, IEEE, 19*(2), 98–100.
Nowak, L. K., & Fox, J. (2018). Avatars and computer mediated communication: A review of the definitions, uses, and effects of degital representations. *Review of Communication Research, 6*, 30–53. https://doi.org/10.12840/issn.2255-4165.2018.06.01.015
Oh, S. Y., Bailenson, J., Krämer, N., & Li, B. (2016). Let the avatar brighten your smile: Effects of enhancing facial expressions in virtual environments. *PLoS One, 11*, Article 161794. https://doi.org/10.1371/journal.pone.0161794
Park, W., Heo, J., & Lee, J. (2021). Talking through the eyes: User experience design for eye gaze redirection in live video conferencing. In *International conference on human-computer InteractionHCII 2021: Human-computer interaction. Interaction techniques and novel applications* (pp. 75–88).
Peper, E., Wilson, V., Martin, M., Rosegard, E., & Harvey, R. (2021). Avoid Zoom fatigue, be present and learn. *NeuroRegulation, 8*(1), 47. https://doi.org/10.15540/nr.8.1.47
Rodeghero, P., Zimmermann, T., Houck, B., & Ford, D. (2021). Please turn your cameras on: Remote onboarding of software developers during a pandemic. In *2021 IEEE/ACM 43th international conference on software engineering: Software engineering in practice track (ICSE-SEIP)*. IEEE, 2021.
Shin, M., Kim, S., & Biocca, F. (2019). The uncanny valley: No need for any further judgements when an avatar looks eerie. *Computers in Human Behavior, 94*, 100–109. https://doi.org/10.1016/j.chb.2019.01.016. . (Accessed 27 May 2019) Accessed.
Spataro, J. (2020). *Remote work trend report: Meetings Microsoft. Microsoft365* https://www.microsoft.com/en-us/microsoft-365/blog/2020/04/09/remote-work-trend-report-meetings/.
Sumioka, H., Yamato, N., Shiomi, M., & Ishiguro, H. (2021). A minimal design of a human infant presence: A case study toward interactive doll therapy for older adults with dementia. *The World for Scientists*, 1–9. https://easychair.org/publications/preprint/Fq7T.
Takano, M., & Taka, F. (2022). Fancy avatar identification and behaviors in the virtual world: Preceding avatar customization and succeeding communication. *Computers in Human Behavior Repots, 6*, Article 100176, 2022.
Tobi, B., Osman, W. H., Bakar, A. L. A., & Othman, I. W. (2021). A case study on students' reason for not switching on their cameras during online class sessions, Special Issue: Supporting the Continuation of Language Teaching and Learning at Malaysian Higher Education Institution. *Learning, 6*(41), 216–224. https://doi.org/10.35631/IJEPC.641016. July 2021.
Walther, J., & D'Addario, K. P. (2001). The impacts of emoticons on message interpretation in computer-mediated communication. *Social Science Computer Review, 19*, 323–345.