

RESEARCH ARTICLE

Can an android's posture and movement discriminate against the ambiguous emotion perceived from its facial expressions?

Satoshi Yagi^{1,2}, Yoshihiro Nakata^{1,2#a*}, Yutaka Nakamura^{1,2#b}, Hiroshi Ishiguro^{1,2}

1 Department of Engineering Science, Osaka University, Toyonaka, Osaka, Japan, **2** JST ERATO, Chiyodaku, Tokyo, Japan

#a Current address: Department of Mechanical and Intelligent Systems Engineering, The University of Electro-Communications, Chofu, Tokyo, Japan

#b Current address: RIKEN Information R&D and Strategy Headquarters, RIKEN, Sorakugun, Kyoto, Japan
* ynakata@uec.ac.jp



Abstract

Expressing emotions through various modalities is a crucial function not only for humans but also for robots. The mapping method from facial expressions to the basic emotions is widely used in research on robot emotional expressions. This method claims that there are specific facial muscle activation patterns for each emotional expression and people can perceive these emotions by reading these patterns. However, recent research on human behavior reveals that some emotional expressions, such as the emotion “intense”, are difficult to judge as positive or negative by just looking at the facial expression alone. Nevertheless, it has not been investigated whether robots can also express ambiguous facial expressions with no clear valence and whether the addition of body expressions can make the facial valence clearer to humans. This paper shows that an ambiguous facial expression of an android can be perceived more clearly by viewers when body postures and movements are added. We conducted three experiments and online surveys among North American residents with 94, 114 and 114 participants, respectively. In Experiment 1, by calculating the entropy, we found that the facial expression “intense” was difficult to judge as positive or negative when they were only shown the facial expression. In Experiments 2 and 3, by analyzing ANOVA, we confirmed that participants were better at judging the facial valence when they were shown the whole body of the android, even though the facial expression was the same as in Experiment 1. These results suggest that facial and body expressions by robots should be designed jointly to achieve better communication with humans. In order to achieve smoother cooperative human-robot interaction, such as education by robots, emotion expressions conveyed through a combination of both the face and the body of the robot is necessary to convey the robot's intentions or desires to humans.

OPEN ACCESS

Citation: Yagi S, Nakata Y, Nakamura Y, Ishiguro H (2021) Can an android's posture and movement discriminate against the ambiguous emotion perceived from its facial expressions? PLoS ONE 16(8): e0254905. <https://doi.org/10.1371/journal.pone.0254905>

Editor: Marina A. Pavlova, University Hospitals Tübingen: Universitätsklinikum Tübingen, GERMANY

Received: January 25, 2021

Accepted: July 6, 2021

Published: August 10, 2021

Copyright: © 2021 Yagi et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All photo and video files are available from the Zenodo database (<https://doi.org/10.5281/zenodo.4983274>).

Funding: This research was partially supported by JST ERATO Grant Number JPMJER1401 and Grant-in-Aid for JSPS Fellows Grant Number JP19J20127.

Competing interests: The authors have declared that no competing interests exist.

Introduction

A number of studies have pointed out that nonverbal communication from robots to humans is as important for robots to communicate smoothly with humans in society as it is in human-

to-human communication [1–4]. Facial expressions and gestures are representative of the nonverbal communication that people often use to convey their emotions [5, 6]. Therefore, it has been actively studied how robots' facial expressions and gestures can convey emotions to achieve meaningful communication with humans [7–9]. For example, humanoids' human-like upper body gestures (nonverbal communication) leads people to perceive a higher animacy of the robot, and the gestures also affect the emotional state and self-disclosure, compared to robot-specific nonverbal behavior (such as LED-eyes color changes) [10]. Also, humanoids' facial expressions lead to an increase in people's desire to interact again with the robot [11]. While there are many findings showing that nonverbal expressions by robots play an important role in human-robot communication, it is also reported that slight differences in expression can lead to the conveyance of significant misinformation [12]. Thus, there is still room for further investigation into the implementation of robotic facial expression, and the human perception of such expressions.

It is widely recognized that facial expressions are important for judging human emotions objectively and clearly. In the 19th century, Darwin and Prodger (1998) already described how facial expressions are associated with certain emotions [13]. Subsequently, Ekman et al. (1987) suggested that facial expressions had universality, and classified six types of basic emotions by different facial expressions [14]. Russell (1980) explained the correspondence of facial expressions and emotions by using two dimensional affective scales: valence (positive to negative) and arousal (high to low) [15]. All of this research focused on mapping facial expressions to basic emotions, assuming that there are specific facial muscle activation patterns for each emotional expression, and that other people can perceive these emotions by reading these patterns. This method of mapping facial expressions to certain emotions has also been used in research that enables robots to communicate their emotions to humans. Breazeal (2003) pioneered research on machines imitating emotional expressions, arguing for the importance of facial expressions and eye gaze for a social robot [16]. Later humanoid robots, such as Kobian, Flobi, Bert2, and iCub, would also use this mapping method, with all four having human-like facial features that allowed for basic emotional expressions [17–20]. The android robots developed in our research group, such as Geminoid, took this one step further, with a realistic appearance created by using silicon-made skin and an original actuation mechanism for the face [21]. This allowed the android to express emotions through facial expressions to an almost human degree [22].

Recent research has found some cases in which emotions may not be clearly distinguishable from facial expressions [23]. Meeren, Heijnsbergen, and Gelder (2005) reported that emotional perception is hindered when facial expression differs from body expression (for example, an angry face, but a frightened body posture or vice versa) [24]. Van den Stock, Righart, and Gelder (2007) also reported that the perception of emotions from facial expressions is strongly influenced by body posture [25]. Aviezer, Trope, and Todorov (2012) found that when participants were shown peak expressive reactions to winning and losing points in professional high-stakes tennis matches, it was difficult for participants to correctly judge whether the emotion was positive or negative (i.e. a win or a loss) through intense facial expressions alone, however, they could perceive emotions when these were seen with the body [26]. Aviezer et al. stated that although the faces were inherently ambiguous, viewers erroneously reported perceiving valence in the face and this process seemed to be automatic as participants had little awareness of the actual facial ambiguity and the original diagnostic source of the valence. In addition, later research also showed that intense emotions cannot be judged solely by facial expressions [27].

It was found that humans recognize facial emotions by unconsciously considering factors other than facial expression. In previous research on robots that have emotional expression functions, facial expressions have been used to convey emotions on the assumption that facial

expressions have a one-to-one correspondence with certain emotions. Thus, it still has not been investigated whether robots can express the ambiguous facial expressions with no clear valence which humans sometimes show, and what these look like. Furthermore, it has also not been investigated whether the robot's emotional valence can be clearly determined from such expressions by adding body expressions.

What we validate in this paper are the following: 1) what kind of facial expression of an android is indistinguishable (ambiguous) in the scale of positive to negative emotions?, and 2) with an indistinguishable facial expression, do postures and movements solve the facial ambiguity? We believe that the investigation of ambiguity expressed by the facial expressions of an android and the clear emotional expressions (positive and negative) achieved by adding body modality can help to realize smooth communication between social robots and humans.

Here we present facial expressions, postures and movements of an android used for experiments. As each component of the android's body can be controlled individually, we conducted three step-by-step experiments and online surveys. In Experiment 1, we validated facial expressions which cannot be clearly discriminated to be either positive or negative, and in Experiment 2 and 3, we validated the possibility of clearly perceiving ambiguous facial expressions of the android by adding postures or movements. In order to investigate only the perception of the peak moment of the android's emotional expressions, we showed posture photos to participants in Experiment 2. Then, in order to investigate the perception of practical emotional expressions, we showed movement movies to participants in Experiment 3. To assess a perceived emotion, we used the emotional state which consists of two dimensions, Valence (Positive—Negative) and Intensity (High—Low), which is often used in psychological research [28].

Materials and methods

We used the child-like mobile android that we developed called *ibuki* [29], who is 120 cm tall and comprised of two parts—a mobility unit (lower body, which contains a linear joint that generates a vertical motion of the upper body) [30] and an upper body, both designed based on the dimensions of a 10-year-old Japanese boy. For a human-like appearance, the face and hands are covered with silicone skin. An electric motor drives each of the 47 joints, and mobile batteries are used as the power supply.

The protocol was approved by the ethics committee for research involving human subjects at the Graduate School of Engineering Science, Osaka University (#R1–6). We recruited separate samples of participants using Amazon Mechanical Turk. The number of participants in each assessment is as follows: 94 people in Experiment 1 (41 females and 53 males, Mean age (M) = 35.5 years old, Standard Deviation (SD) = 10.72), 114 people in Experiment 2 (44 females and 70 males, M = 34.9, SD = 9.93), and 114 people in Experiment 3 (53 females and 61 males, M = 35.4, SD = 9.92).

In Experiment 1, participants were asked to answer the following two questions, corresponding to the emotional state consisting of two dimensions: valence and intensity, after viewing *ibuki's* facial expressions:

1. To rate the robot's facial emotion (negative to positive valence) from a scale of -4 to 4.
2. To rate the intensity of the robot's facial expression from a scale of 1 to 9.

In Experiment 2 and 3, participants were asked to watch a photo or movie that showed the moment *ibuki* reacted to the result of a game and instructed to guess whether *ibuki* had won or lost the game based on its facial expression. In addition to the questions from Experiment 1, two more questions were added (interpretation of the game result and human-likeness):

1. To guess whether they think the robot has won or lost the game.
2. To rate the robot's facial emotion (negative to positive valence) from a scale of -4 to 4.
3. To rate the intensity of the robot's facial expression from a scale of 1 to 9.
4. To rate the human-likeness of the robot's facial expression from a scale of 1 to 9.

The question about human-likeness was added to confirm that the human likeness of *ibuki* did not affect the assessment of the valence, since the postures and movements were manually created by the authors.

Procedure

In Experiment 1, we investigated whether there are ambiguous facial expressions which cannot be determined by the facial expression alone. For facial expressions, we created nine facial expressions, namely: anger, disgust, fear, happiness, sadness, surprise, contempt, neutral, which are representative emotions in facial expression studies, and intense. For facial expressions except intense, we operated actuators with reference to the Emotional Facial Action Coding System, which explains the characteristics of actual human facial movements. For intense, we created the intense facial expression as an expression of high muscle activity that expresses the excitement immediately after a game result was decided. It was created to look like an emotion in which the eyes are closed strongly and the mouth expresses shouting (See Fig 1) based on the paper of Aviezer et al. [26].

From those eight expressions except neutral, we selected one facial expression with the highest entropy of the valence in order to use it in Experiment 2 and 3. Since entropy is a physical property that represents a state of disorder of the system, for this study, we defined the indistinguishable (ambiguous) facial expression as having the highest entropy of the facial valence distributions. The entropy S of each facial expression is calculated by the following equation:

$$S = - \sum_{x_i=-4}^4 P(x_i) \log_2 P(x_i) \quad (1)$$

where $x_i = \{-4, -3, \dots, 3, 4\}$ is the possible facial valence, $P(x_i)$ is the probability which the facial valence is rated as x_i . The calculation method is further detailed in S1 Appendix. In Experiment 1, one photo of the eight emotions (350 x 450 pixels size) was displayed beside the neutral face on each page in a random order. After each viewing, participants were asked to answer the facial valence and intensity by answering the two questions mentioned in the Materials and methods section, taking into account that the valence and intensity of the neutral face was 0 and 1, respectively.

Together with the highest entropy facial expression in Experiment 1, we took photos of 15 types of postures by *ibuki* for Experiment 2. In addition, we took movies of 30 types of movements performed by *ibuki*. Firstly, we constructed three arm poses (AP; labeled as A, B, and C) and five head angles (HA; labeled as -43, -17, 0, 17, and 43, which are head inclination degrees at the sagittal plane). HA consist of the neck joint angle θ_1 and waist joint angle θ_2 as shown in Fig 2. Table 1 shows the angles of the head, neck, and waist.

Secondly, we created 30 movements by combining the previous 15 postures and two types of vertical motions (VM) of the upper body: one was a +40 mm upward motion, the other was a -40 mm downward motion (VM; labeled as 40 and -40). Then, we placed two cameras—in front of and 40 degrees diagonally to *ibuki*—to take photos and movies of *ibuki*'s emotional expressions. Both cameras were installed at a height of 100 cm—assuming the position of an adult on the knees looking at a child 140 cm away (Fig 3). Behind *ibuki* was a neutral green

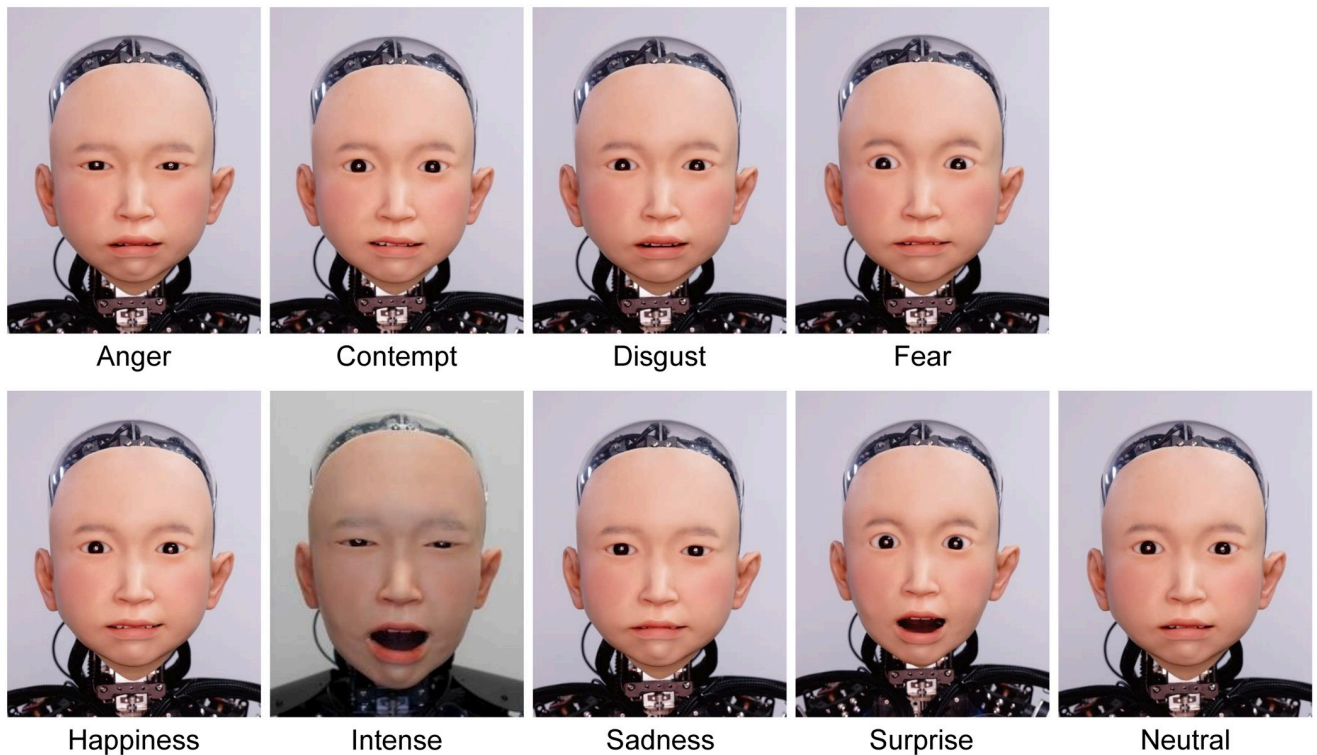


Fig 1. Nine facial expressions of *ibuki*.

<https://doi.org/10.1371/journal.pone.0254905.g001>

screen. For Experiment 2, we took two photos of each of *ibuki*'s postures from the front and diagonal angles, resulting in 30 photos. Fig 4 shows 15 postures from the front view. For Experiment 3, we took two movies of each movement from the front and diagonal angles, resulting in 60 movies. Each movie was edited to a length of three seconds. The first 0.5 seconds of the movie showed *ibuki* in the neutral posture before moving towards the target posture. Again, the same fixed facial expression was used for both Experiment 2 and 3.

In Experiment 2 and 3, we investigated whether the indistinguishable facial expression can be distinguished by adding body postures and movements. Participants were asked to look at the photos or movies. To eliminate the influence on the assessment of valence due to viewpoint changes, one participant group (57 people) only saw the front angle photos, and the other participant group (57 people) looked at photos taken from the diagonal angle. In order for participants to understand the contents of the questions, we used a context that evaluate the reaction of the child android at the moment when the win or loss of a game was decided. At the beginning of each assessment, the neutral upright posture of *ibuki* was shown as a reference. For both experiments, a photo or movie was displayed on each page in random order one by one. After each viewing, participants were asked to answer the aforementioned four questions (interpretation of the game result, facial valence, intensity and human-likeness).

For the processing of the data of each experiment, only participants who answered the manipulation check questions correctly were counted as valid answers (this question asked participants to answer with a certain score as was instructed in the movie to check if the participants properly watched the movies). In addition, in Experiment 2, the neutral posture with the neutral face that was shown at first as a reference was shown twice during the task, and participants who rated the valence of the neutral posture as -4 or 4 had their answers excluded

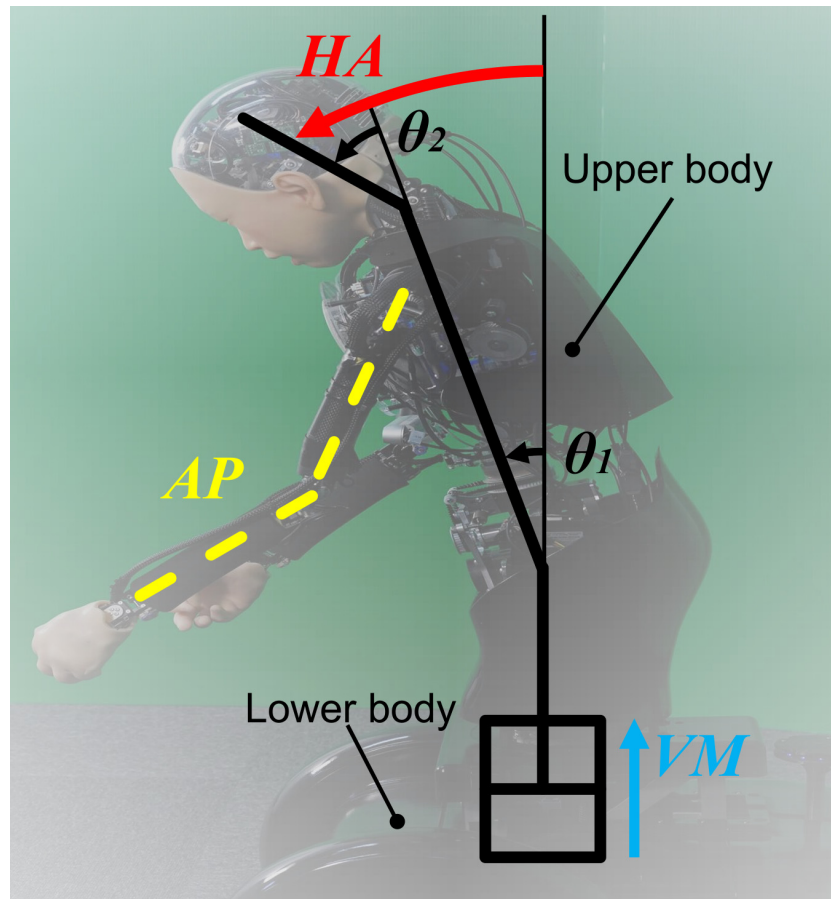


Fig 2. The body expression with the ambiguous facial expression and the configuration of HA, AP, and VM.

<https://doi.org/10.1371/journal.pone.0254905.g002>

(resulting in 10 exclusions). In Experiment 3, the movie of condition HA/AP/VM: -17/A/40 was shown three times during the task, and participants who provided answers with a more than five points difference were excluded (resulting in 23 exclusions).

Analysis

In Experiment 1, distributions of both valence and intensity were analyzed to check the characteristics of *ibuki's* facial expressions. Then we calculated the entropy for the facial valences of each expression to verify the indistinguishable facial expressions, which have a high entropy.

In Experiment 2, the mean facial valences and entropy were analyzed to check the influence of android body expressions. Two-way repeated measures analyses of variance (ANOVA)

Table 1. The relationship among head angle, the neck, and waist joint.

Head angle HA	-43	-17	0	17	43
Neck angle θ_1	-30	-30	0	30	30
Waist angle θ_2	-13	13	0	-13	13

All angles are shown in degrees.

<https://doi.org/10.1371/journal.pone.0254905.t001>

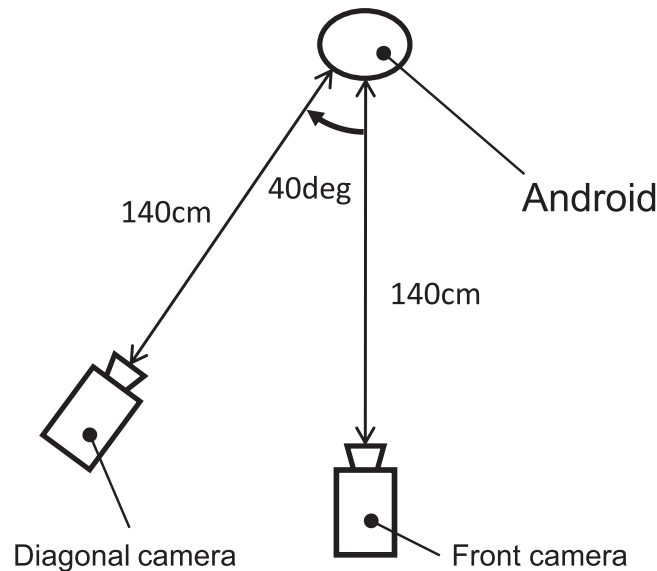


Fig 3. An overview of shooting environment.

<https://doi.org/10.1371/journal.pone.0254905.g003>

were conducted with HA and AP as two within-subjects factors. The significance level was set at 0.05. Partial eta-squared (η_p^2) was reported to demonstrate the effect size in ANOVA. Then Tukey's HSD test (HSD test) was performed for multiple comparisons to verify our hypothesis that adding postures or movements contributes to the clear perception of the ambiguous facial expression. In addition, the percentage that assessed as *won* at each posture was also calculated to verify the distinguishable face due to body expressions.

As an android can control each component of the body individually—e.g. facial expression, head angles and arm poses—instead of in conjunction with other joints and other parts of the body like humans, the postures and movements of *ibuki* cannot be guaranteed to be intense and human-like. Therefore, we calculated correlations between facial valence with intensity and human-likeness to confirm whether the effect of intensity and human-likeness on the assessment of facial valence was small or not.

In Experiment 3, we analyzed the data following the same analysis procedure used in Experiment 2.

Results

Experiment 1

Fig 5 shows the distribution of assessed facial valence for eight emotions in Experiment 1. The horizontal axis shows the assessed facial valence (-4 to 4) and the vertical axis shows the normalized number of responses. The dotted line of the histogram shows the mean facial valence. In order to select the most ambiguous face, we calculated the entropy of each distribution, which is shown in the upper left of the histogram. The highest was intense (3.032). With this result, we decided to use intense as the ambiguous facial expression for the next experiments.

Fig 6 shows the distribution of assessed facial intensity for the eight emotions in Experiment 1. As in the previous valence result (Fig 5), the horizontal axis shows the assessed facial

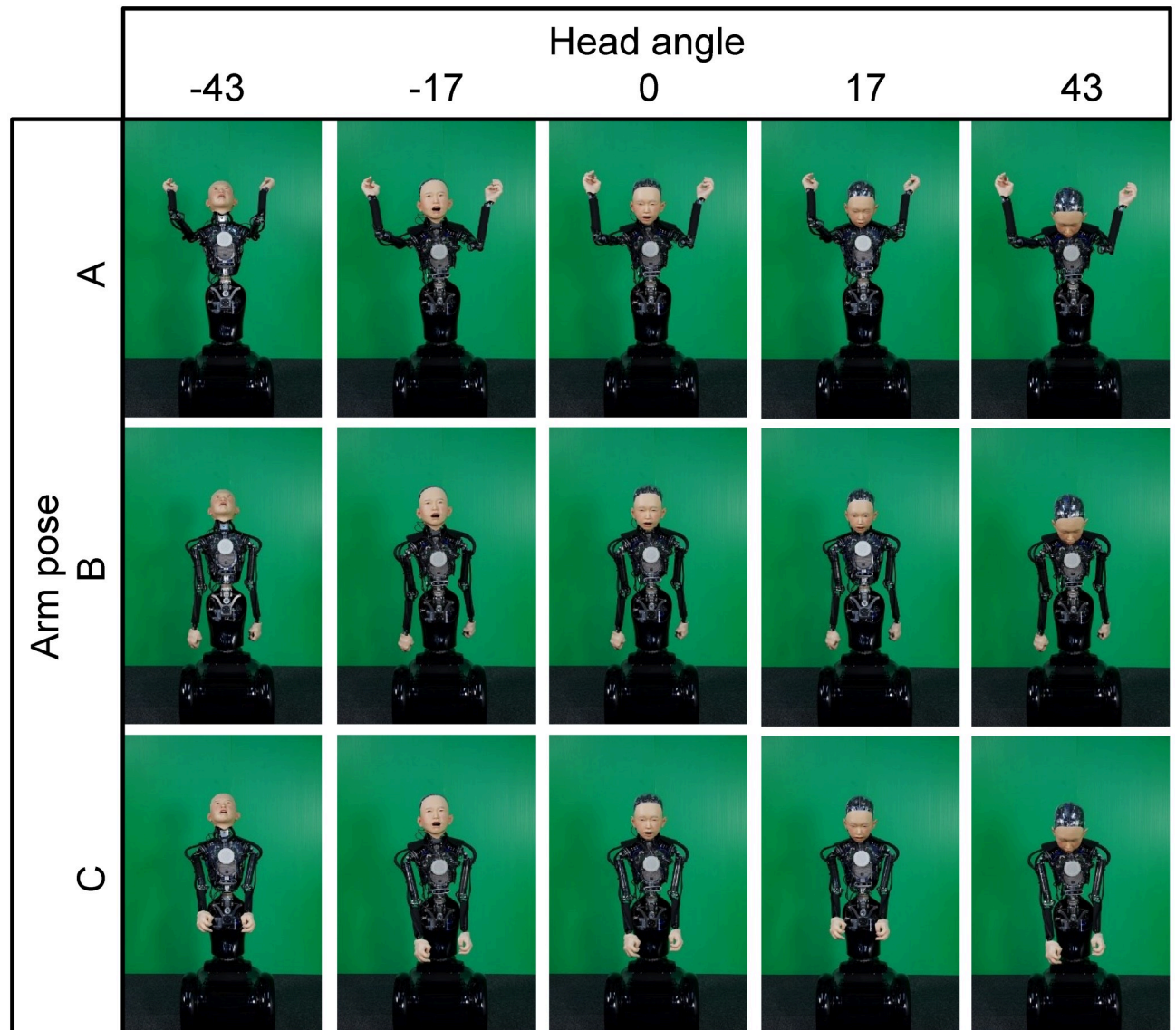


Fig 4. 15 postures from the front view.

<https://doi.org/10.1371/journal.pone.0254905.g004>

intensity (1 to 9) and the vertical axis shows the normalized number of responses. The dotted line in the histogram shows the mean facial intensity.

Experiment 2

In Experiment 2, *ibuki's* posture photos were shown to participants in order to investigate only the perception of the peak moment of emotional expressions. We validated that postures contribute to the perception of the indistinguishable facial expression. Fig 7 shows the distribution of assessed facial valence in Experiment 2. The horizontal axis shows the facial valence as assessed by participants and the vertical axis shows the normalized number of answers for each posture. The red line histogram shows the result of the intense facial expressions in

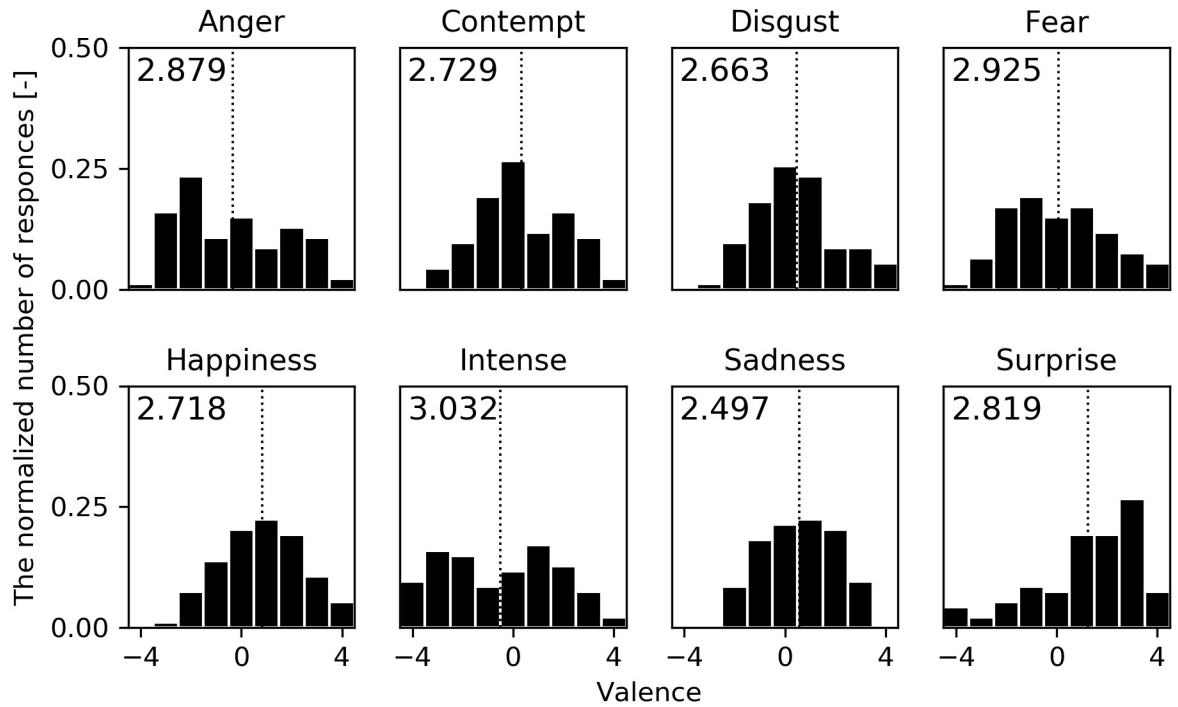


Fig 5. Distributions of the assessed facial valence in Experiment 1. The horizontal axis shows the assessed facial valence (-4 to 4) and the vertical axis shows the normalized number of responses. The dotted line shows the mean facial valence. The entropy is shown in the upper left of each histogram.

<https://doi.org/10.1371/journal.pone.0254905.g005>

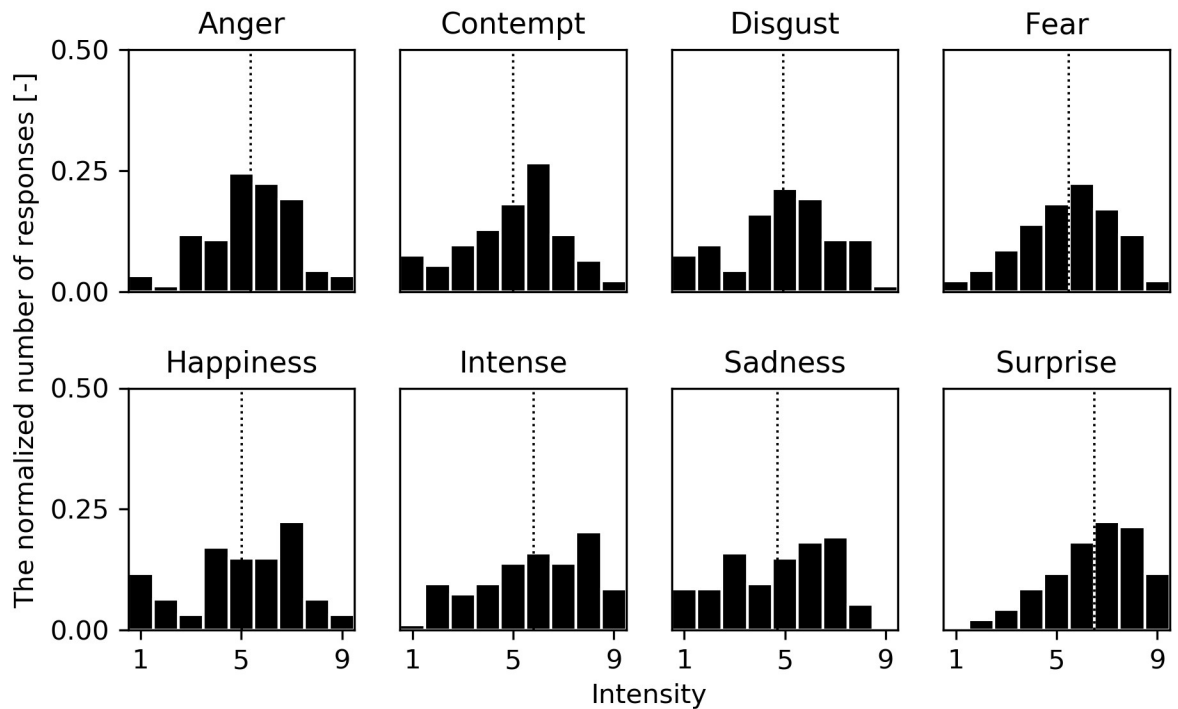


Fig 6. Distributions of the assessed facial intensity in Experiment 1. The horizontal axis shows the assessed facial intensity (1 to 9) and the vertical axis shows the normalized number of responses. The dotted line shows the mean facial intensity.

<https://doi.org/10.1371/journal.pone.0254905.g006>

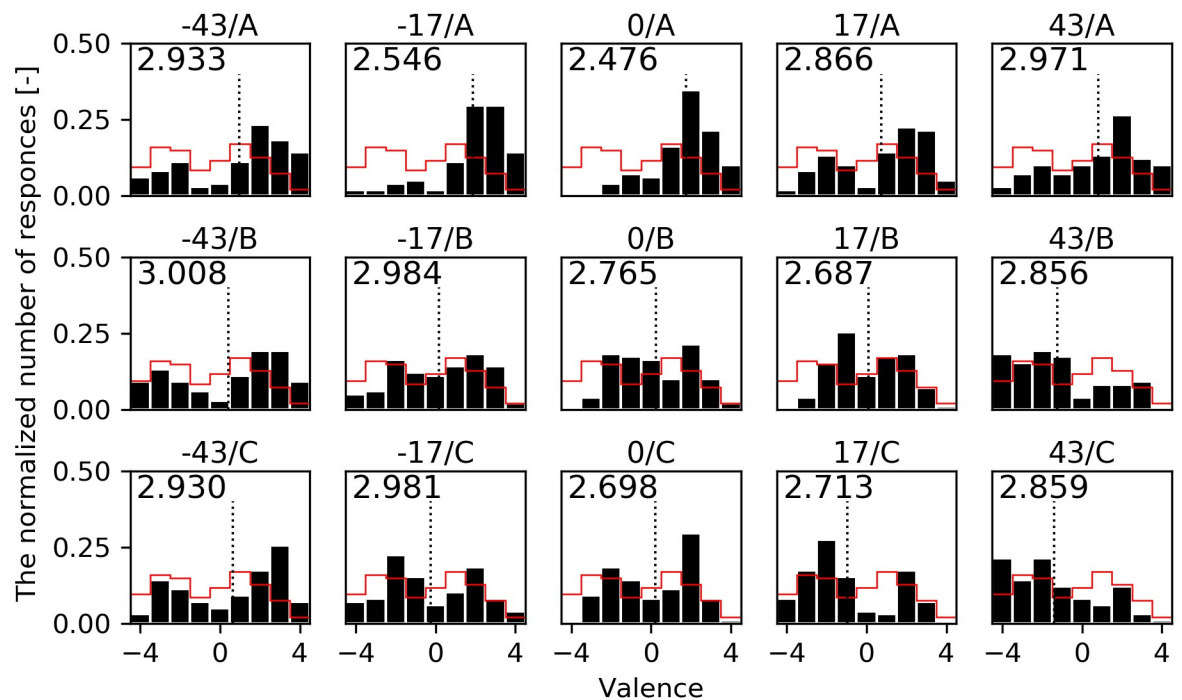


Fig 7. Distributions of the assessed facial valence in Experiment 2. The horizontal axis shows the assessed facial valence (-4 to 4) and the vertical axis shows the normalized number of responses. The dotted line shows the mean facial valence. The red line on the histogram shows the result of the intense facial expressions in Experiment 1 as a reference. The entropy is shown in the upper left of each histogram.

<https://doi.org/10.1371/journal.pone.0254905.g007>

Experiment 1 as a reference. We calculated the entropy of each distribution, which is shown in the upper left of the histogram. The highest was -43/B and the lowest was 0/A.

In order to investigate whether the mean values difference of facial valences were affected by HA and AP, we ran a two-way repeated measures ANOVA with HA and AP. For the assessment of facial valence, there were significant main effects of HA ($F(4, 113) = 23.86, p < 0.001, \eta_p^2 = 0.034$) and AP ($F(2, 113) = 86.16, p < 0.001, \eta_p^2 = 0.202$) and there was also a significant interaction effect of the HA/AP ($F(8, 113) = 9.89, p < 0.001, \eta_p^2 = 0.055$). Then, we ran HSD test to verify the difference of the mean values for the 15 postures. Table B in [S1 Appendix](#) shows the mean differences for each posture. The highest valence (most positively perceived) posture was HA: -17 degrees and AP: A ([Fig 8a](#), Mean value = 1.89, Standard Error = 0.19), while the lowest valence (most negatively perceived) posture was HA: 43 degrees and AP: C ([Fig 8b](#), $M = -1.40, SE = 0.22$), with a significant difference of 3.29 between these two postures ($p < 0.001$).

[Table 2](#) shows the percentages of people that assessed each posture as *won* in Experiment 2. AP: A was assessed as *won* between 54.1% and 83.7%, in contrast AP: B and AP: C were assessed as *lost* (lower than 50%). The result is also consistent with the assessment tendency of the facial valence.

Looking at HA: -17, the maximum value (83.7%) was under AP: A, on the contrary, the same HA: -17 was also the minimum value (27.6%) under AP: C. Interestingly, there was a 56.1 points gap between AP: A and AP: C even with the same HA.

[Figs A and B in S1 Appendix](#) show the distribution of facial intensity and human-likeness in Experiment 2. The total average of intensity was 6.14 with $SE = 0.04$ and human-likeness

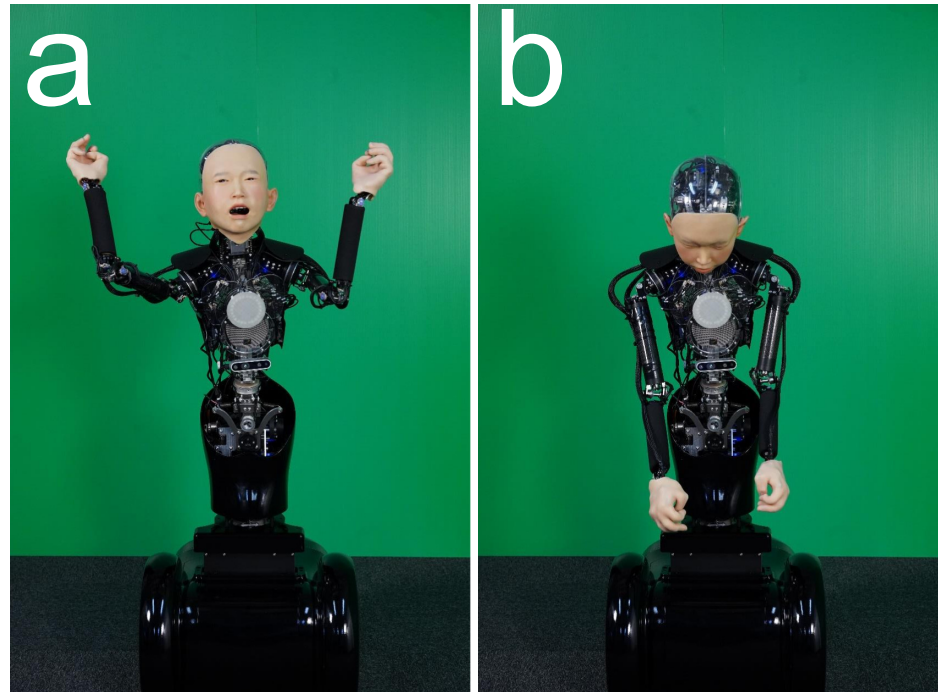


Fig 8. Two representative postures which have distinguishable facial expressions by adding body expressions. (a) The posture -17/A at the highest facial valence / (b) The posture 43/C at the lowest facial valence.

<https://doi.org/10.1371/journal.pone.0254905.g008>

was 6.43 with SE = 0.04. Furthermore, the facial valence had a low correlation with intensity ($r = 0.163, p < 0.001$) and with human-likeness ($r = 0.138, p < 0.001$).

Experiment 3

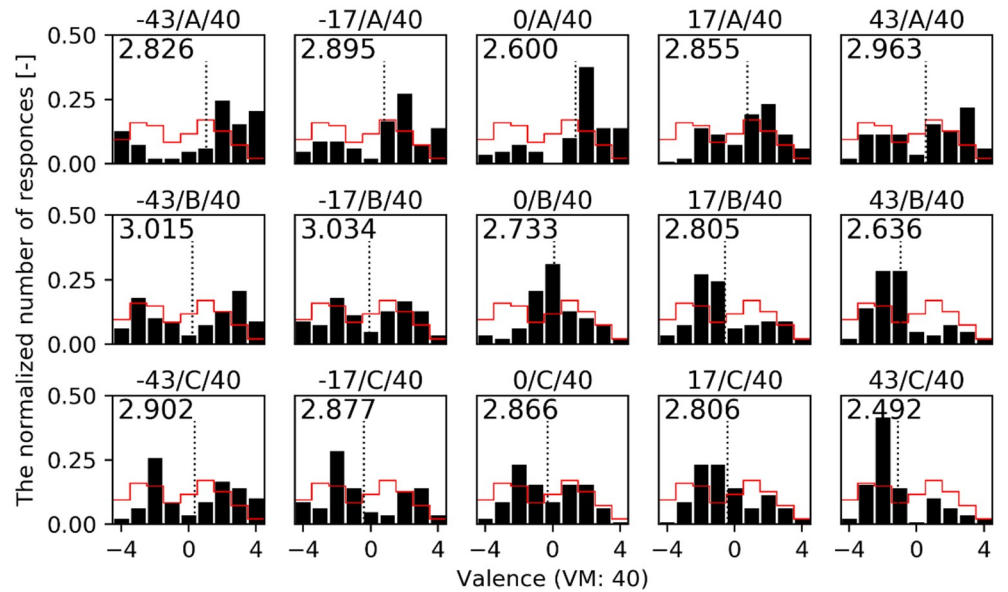
In Experiment 3, *ibuki's* movement movies were shown to participants in order to investigate the perception of emotional expressions. We validated that movements contribute to the perception of the indistinguishable facial expression. Fig 9 shows the distribution of assessed facial valence for the condition VM: 40 and -40 mm in Experiment 3. The horizontal axis shows the facial valence as assessed by participants and the vertical axis shows the normalized number of answers for each movement. The red line histogram shows the result of intense facial expression in Experiment 1 as a reference. We calculated the entropy of each distribution, which is shown in the upper left of the histogram. The highest was -43/C/-40 and the lowest was 43/A/40.

In order to investigate whether the mean values difference of facial valence were affected by HA, AP, and VM, we ran a three-way repeated measures ANOVA with HA, AP, and VM. For

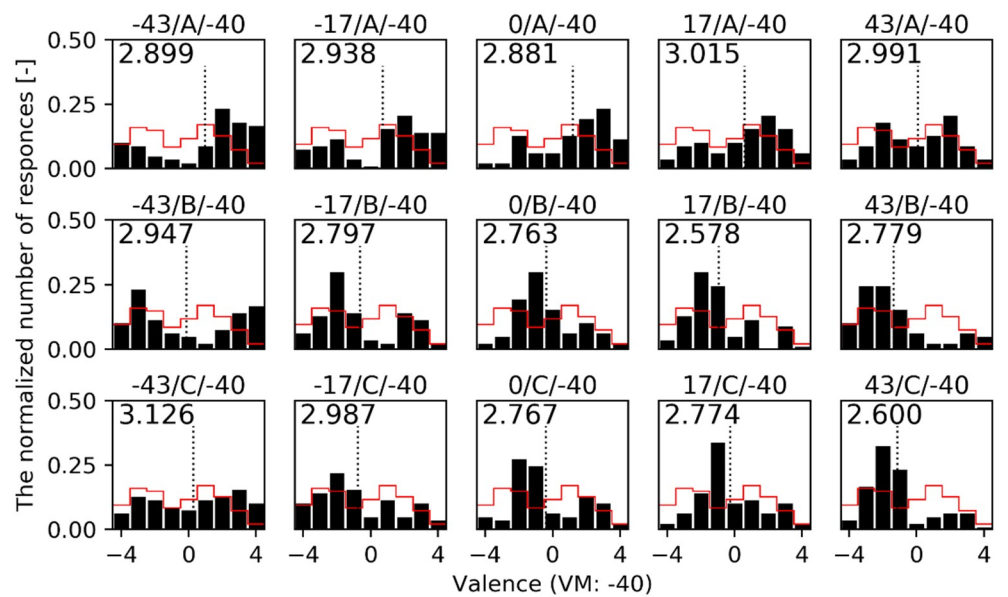
Table 2. Percentage of participants that judged each posture as won a game in Experiment 2.

		HA				
		-43	-17	0	17	43
AP	A	61.2	83.7	82.7	54.1	55.1
	B	50.0	36.7	35.7	25.5	5.1
	C	49.0	27.6	42.9	6.1	7.1

<https://doi.org/10.1371/journal.pone.0254905.t002>



(a) condition VM: 40 mm



(b) condition VM: -40 mm

Fig 9. Distributions of the assessed facial valence in Experiment 3. The horizontal axis shows the facial valence as assessed by participants and the vertical axis shows the normalized number of answers at each movement. The dotted line shows the mean facial valence. The red line on the histogram shows the result of intense facial expression in Experiment 1 as a reference. The entropy is shown in the upper left of each histogram.

<https://doi.org/10.1371/journal.pone.0254905.g009>

Table 3. Percentage of participants that judged each movement as won a game in Experiment 3.

		HA				
		-43	-17	0	17	43
AP/VM	A/40	69.7	64.5	76.3	63.2	50.0
	B/40	47.4	43.4	46.1	23.7	11.9
	C/40	52.6	28.9	40.8	23.7	18.4
	A/-40	68.4	63.2	75.0	65.8	43.4
	B/-40	39.5	30.3	25.0	14.5	17.1
	C/-40	51.3	31.6	30.3	30.3	13.2

<https://doi.org/10.1371/journal.pone.0254905.t003>

the assessment of facial valence, there were significant main effects for all three: HA ($F(4, 113) = 11.68, p < 0.001, \eta_p^2 = 0.020$), AP ($F(2, 113) = 52.82, p < 0.001, \eta_p^2 = 0.044$), and VM ($F(1, 113) = 17.35, p < 0.001, \eta_p^2 = 0.007$). Also there were significant interaction effects of the HA/AP ($F(8, 113) = 3.29, p = 0.001, \eta_p^2 = 0.011$) and the AP/VM ($F(2, 113) = 3.99, p = 0.021, \eta_p^2 = 0.003$). On the other hand, no significance was found in the HA/VM ($F(4, 113) = 0.48, p = 0.748, \eta_p^2 = 0.001$) or the HA/AP/VM ($F(8, 113) = 0.57, p = 0.806, \eta_p^2 = 0.002$).

Thus, we ran Tukey's HSD tests to verify the difference of the mean values at each interaction effect of HA/AP and AP/VM. The result of HA/AP is shown in Table C in [S1 Appendix](#) and AP/VM is shown in Table D in [S1 Appendix](#). Under the condition of HA and AP, the highest valence movement was HA: 0 degrees and AP: A ($M = 1.28, SE = 0.18$), while the lowest valence movement was HA: 43 degrees and AP: B ($M = -1.14, SE = 0.17$), with a significant difference of 2.42 between these two movements ($p < 0.001$). Also under the condition of AP and VM, the highest valence movement was AP: A and VM: 40 mm ($M = 0.90, SE = 0.12$), while the lowest valence movement was AP: B and VM: -40 mm ($M = -0.69, SE = 0.12$), with a significant difference of 1.59 between these two movements ($p < 0.001$).

[Table 3](#) shows the percentages of people that assessed each movement as *won* in Experiment 3.

As in Experiment 2, different AP leads to the exact opposite judgement for the game result. For example, under HA: 17, there is a 39.5 points gap between A/40 (63.2%) and B/40 (23.7%) and a 51.3 points gap between A/-40 (65.8%) and B/-40 (14.5%).

Figs C and D in [S1 Appendix](#) show the distribution of facial intensity and human-likeness in Experiment 3. The total average of intensity was 6.00 with $SE = 0.04$ and human-likeness was 6.02 with $SE = 0.04$. Furthermore, the facial valence had a low correlation with intensity ($r = 0.164, p < 0.001$) and with human-likeness ($r = 0.196, p < 0.001$).

Discussion

In the past, research on robots expressing emotions has been based on the assumption that a certain expression is equivalent to a certain emotion (e.g. a smiling indicates happiness). However, there has been no investigation into whether robots can also express ambiguous facial expressions with no clear valence. Additionally, for ambiguous facial expressions, there has been no investigation into whether the addition of expressions such as body expression can make the robot's emotional valence clearer to humans.

In Experiment 1, we validated that the distributions of valence and intensity were different depending on the facial emotion expressions. In particular, according to the results of the entropy of facial valences, participants could not distinguish the facial valence of the intense emotion as positive or negative. The previous study by Aviezer et al. found that it is difficult to

judge the valence of intense facial expressions when only the facial expression can be seen [26]. Our experiment corroborates this phenomenon in the case of androids. The histogram of the facial valence of intense emotion in Fig 5 seems to have two peaks in both sides positive (1) and negative (-3). We believe that this is because the participants were mainly divided into two groups who perceived the intense facial expression positively or negatively. Thus, the entropy of intense got the highest value of the eight emotions.

In Experiment 2 and 3, it was found that both body postures and body movements changed the facial valence of the intense emotion and contributed to the lower entropy in the valence assessments. In other words, postures and movements can improve the perception of the ambiguous facial expression. As mentioned in the result, the maximum difference in the mean facial valences was 3.29 for the posture and 2.42 for the movement. Although the number of facial expressions and postures used in the experiment is different, as a reference, those mean differences are as close as the mean difference of 2.3 points in the previous human study by Aviezer et al., in which they exchanged win and lose body postures against intense face in a tennis match. Looking at the histograms of facial emotional values in Fig 7 of Experiment 2, under postures which had lower entropy than that of the intense facial expression in Experiment 1, such as -17/A (0.486 lower) and 0/A (0.556 lower), those distributions only had a one side peak. On the other hand, under postures which had little lower entropy, such as -43/B and -17/B, those distributions still had two peaks in both positive and negative sides as in Experiment 1. However, -17/B/-40 (the movement of -17/B with VM: -40) tended to be perceived more negatively due to the effect of the downward vertical motion in Experiment 3. Thus, -17/B/-40 had a further lower entropy than that of Experiment 2.

From the η_p^2 values in ANOVA, we can deduce that participants' assessments were most affected by AP, then HA and, finally, by VM. Our findings are consistent with several previous studies on human emotion expressions. As for AP, when the arms are located vertically upwards or move upwards from the body, the posture tends to be perceived as joy [31, 32] and joy corresponds to positive valence [15]. As for HA, raised head angles are associated with positive emotions such as joy [33]. As for VM, rising upper body motion corresponds to happiness, and sinking body motion corresponds to sadness [34]. However, in this study, we only controlled the displacement and not the speed and acceleration of the movement. It has been found that the human perception of robots is negatively impacted if human-like robots perform robot-like movements which contradict its appearance [35]. This might be the reason why VM comes last. Motion activity, including how fast and smooth a person or android is moving, is said to be one of the important factors for emotion recognition [36].

Since interaction effects were confirmed between AP and HA, and AP and VM from the results of ANOVA, we argue that robot developers should not separately design facial expressions from body expressions when designing robots for the application of human-robot interaction, as there are expressions that cannot be distinguished by the facial expressions alone, such as the intense emotion discussed in this paper. Strong emotions influence human social decision making [37], enhancement on memory [38], and time perception [39]. In the same way, strong emotion expressions by robots are expected to enhance the relationship between humans and robots [40].

One limitation of this study was that *ibuki* was based on a young Asian male child. In past studies on recognition of human emotions, it has been reported that age, gender, and ethnicity can influence emotional perception [41, 42]. Therefore, it is necessary to further investigate how our finding applies to other common humanoids and robots universally. Previous research on humans shows that humans can perceive emotions even from point-light displayed facial expressions or body movements [5, 43]. We expect that this capability is also applicable for the emotion perception of robot faces.

Conclusion

In this paper, we conducted three experiments by controlling the android face and body components. In Experiment 1, we validated a facial expression with no clear valence. Based on the entropy of the valence distribution, the intense facial expression was the most ambiguous. Then we validated whether the addition of postures or movements could make the android's facial valence clearer to humans in Experiment 2 and 3. We confirmed the possibility of clearly perceiving the android's facial expression of intense by adding postures or movements.

Supporting information

S1 Appendix. The summary of the experiment result.
(PDF)

Author Contributions

Conceptualization: Satoshi Yagi, Yoshihiro Nakata, Yutaka Nakamura.

Data curation: Satoshi Yagi.

Formal analysis: Satoshi Yagi.

Funding acquisition: Hiroshi Ishiguro.

Investigation: Satoshi Yagi.

Methodology: Satoshi Yagi.

Supervision: Yoshihiro Nakata, Yutaka Nakamura, Hiroshi Ishiguro.

Writing – original draft: Satoshi Yagi.

Writing – review & editing: Satoshi Yagi, Yoshihiro Nakata, Yutaka Nakamura.

References

1. Adams RB Jr, Kleck RE. Effects of direct and averted gaze on the perception of facially communicated emotion. *Emotion*. 2005; 5(1):3. <https://doi.org/10.1037/1528-3542.5.1.3>
2. Mutlu B, Kanda T, Forlizzi J, Hodgins J, Ishiguro H. Conversational gaze mechanisms for humanlike robots. *ACM Transactions on Interactive Intelligent Systems (TiiS)*. 2012; 1(2):1–33. <https://doi.org/10.1145/2070719.2070725>
3. Sheikholeslami S, Moon A, Croft EA. Cooperative gestures for industry: Exploring the efficacy of robot hand configurations in expression of instructional gestures for human–robot interaction. *The International Journal of Robotics Research*. 2017; 36(5-7):699–720. <https://doi.org/10.1177/0278364917709941>
4. Thepsonthorn C, Ogawa Ki, Miyake Y. The relationship between robot's nonverbal behaviour and human's likability based on human's personality. *Scientific reports*. 2018; 8(1):1–11.
5. Bidet-Ildei C, Decatoire A, Gil S. Recognition of Emotions From Facial Point-Light Displays. *Frontiers in Psychology*. 2020; 11:1062. <https://doi.org/10.3389/fpsyg.2020.01062>
6. Atkinson AP, Dittrich WH, Gemmell AJ, Young AW. Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception*. 2004; 33(6):717–746. <https://doi.org/10.1068/p5096>
7. Trovato G, Kishi T, Endo N, Zecca M, Hashimoto K, Takanishi A. Cross-cultural perspectives on emotion expressive humanoid robotic head: recognition of facial expressions and symbols. *International Journal of Social Robotics*. 2013; 5(4):515–527. <https://doi.org/10.1007/s12369-013-0213-z>
8. Erden MS. Emotional postures for the humanoid-robot nao. *International Journal of Social Robotics*. 2013; 5(4):441–456. <https://doi.org/10.1007/s12369-013-0200-4>
9. Deshmukh A, Craenen B, Vinciarelli A, Foster ME. Shaping Robot Gestures to Shape Users' Perception: The Effect of Amplitude and Speed on Godspeed Ratings. In: *Proceedings of the 6th International Conference on Human-Agent Interaction*; 2018. p. 293–300.

10. Rosenthal-von der Pütten AM, Krämer NC, Herrmann J. The effects of humanlike and robot-specific affective nonverbal behavior on perception, emotion, and behavior. *International Journal of Social Robotics*. 2018; 10(5):569–582. <https://doi.org/10.1007/s12369-018-0466-7>
11. Kedzierski J, Muszyński R, Zoll C, Oleksy A, Frontkiewicz M. EMYS—emotive head of a social robot. *International Journal of Social Robotics*. 2013; 5(2):237–249. <https://doi.org/10.1007/s12369-013-0183-1>
12. Tsiourti C, Weiss A, Wac K, Vincze M. Multimodal integration of emotional signals from voice, body, and context: effects of (in) congruence on emotion recognition and attitudes towards robots. *International Journal of Social Robotics*. 2019; 11(4):555–573. <https://doi.org/10.1007/s12369-019-00524-z>
13. Darwin C, Prodger P. *The expression of the emotions in man and animals*. Oxford University Press, USA; 1998.
14. Ekman P, Friesen WV, O'sullivan M, Chan A, Diacoyanni-Tarlatzis I, Heider K, et al. Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of personality and social psychology*. 1987; 53(4):712. <https://doi.org/10.1037/0022-3514.53.4.712> PMID: 3681648
15. Russell JA. A circumplex model of affect. *Journal of personality and social psychology*. 1980; 39(6):1161. <https://doi.org/10.1037/h0077714>
16. Breazeal C. Emotion and sociable humanoid robots. *International journal of human-computer studies*. 2003; 59(1-2):119–155. [https://doi.org/10.1016/S1071-5819\(03\)00018-1](https://doi.org/10.1016/S1071-5819(03)00018-1)
17. Zecca M, Mizoguchi Y, Endo K, Iida F, Kawabata Y, Endo N, et al. Whole body emotion expressions for KOBAN humanoid robot—preliminary experiments with different Emotional patterns—. In: RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication. IEEE; 2009. p. 381–386.
18. Lütkebohle I, Hegel F, Schulz S, Hackel M, Wrede B, Wachsmuth S, et al. The bielefeld anthropomorphic robot head “Flobi”. In: 2010 IEEE International Conference on Robotics and Automation. IEEE; 2010. p. 3384–3391.
19. Bazo D, Vaidyanathan R, Lentz A, Melhuish C. Design and testing of a hybrid expressive face for a humanoid robot. In: 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE; 2010. p. 5317–5322.
20. Churamani N, Cruz F, Griffiths S, Barros P. iCub: learning emotion expressions using human reward. arXiv preprint arXiv:200313483.
21. Ishiguro H. Android science. In: *Robotics Research*. Springer; 2007. p. 118–127.
22. Becker-Asano C, Ishiguro H. Evaluating facial displays of emotion for the android robot Geminoid F. In: 2011 IEEE workshop on affective computational intelligence (WACI). IEEE; 2011. p. 1–8.
23. Bartneck C, Reichenbach J. Subtle emotional expressions of synthetic characters. *International journal of human-computer studies*. 2005; 62(2):179–192. <https://doi.org/10.1016/j.ijhcs.2004.11.006>
24. Meeren HK, van Heijnsbergen CC, de Gelder B. Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences*. 2005; 102(45):16518–16523. <https://doi.org/10.1073/pnas.0507650102>
25. Van den Stock J, Righart R, De Gelder B. Body expressions influence recognition of emotions in the face and voice. *Emotion*. 2007; 7(3):487. <https://doi.org/10.1037/1528-3542.7.3.487>
26. Aviezer H, Trope Y, Todorov A. Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science*. 2012; 338(6111):1225–1229. <https://doi.org/10.1126/science.1224313>
27. Witkower Z, Tracy JL. Bodily communication of emotion: evidence for extrafacial behavioral expressions and available coding systems. *Emotion Review*. 2019; 11(2):184–193. <https://doi.org/10.1177/1754073917749880>
28. Anderson AK, Christoff K, Stappen I, Panitz D, Ghahremani D, Glover G, et al. Dissociated neural representations of intensity and valence in human olfaction. *Nature neuroscience*. 2003; 6(2):196–202. <https://doi.org/10.1038/nn1001> PMID: 12536208
29. JST ERATO Ishiguro Symbiotic Human–Robot Interaction Project. Improvement of humanlike conversations in humanoid robots—Development of a hild-like android with the ability to move—; 2018. Available from: https://www.jst.go.jp/pr/announce/20180731-2/index_e.html.
30. Yagi S, Ise N, Yu S, Nakata Y, Nakamura Y, Ishiguro H. Perception of Emotional Gait-like Motion of Mobile Humanoid Robot Using Vertical Oscillation. In: *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction. HRI'20*. New York, NY, USA: Association for Computing Machinery; 2020. p. 529–531. Available from: <https://doi.org/10.1145/3371382.3378319>.
31. De Meijer M. The contribution of general features of body movement to the attribution of emotions. *Journal of Nonverbal behavior*. 1989; 13(4):247–268. <https://doi.org/10.1007/BF00990296>

32. Gross MM, Crane EA, Fredrickson BL. Methodology for assessing bodily expression of emotion. *Journal of Nonverbal Behavior*. 2010; 34(4):223–248. <https://doi.org/10.1007/s10919-010-0094-x>
33. Mignault A, Chaudhuri A. The many faces of a neutral face: Head tilt and perception of dominance and emotion. *Journal of nonverbal behavior*. 2003; 27(2):111–132. <https://doi.org/10.1023/A:1023914509763>
34. Shafir T, Tsachor RP, Welch KB. Emotion regulation through movement: unique sets of movement characteristics are associated with and enhance basic emotions. *Frontiers in psychology*. 2016; 6:2030.
35. Saygin AP, Chaminade T, Ishiguro H, Driver J, Frith C. The thing that should not be: predictive coding and the uncanny valley in perceiving human and humanoid robot actions. *Social cognitive and affective neuroscience*. 2012; 7(4):413–422. <https://doi.org/10.1093/scan/nsr025>
36. Glowinski D, Dael N, Camurri A, Volpe G, Mortillaro M, Scherer K. Toward a minimal representation of affective gestures. *IEEE Transactions on Affective Computing*. 2011; 2(2):106–118. <https://doi.org/10.1109/T-AFFC.2011.7>
37. Winston JS, Strange BA, O'Doherty J, Dolan RJ. Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature neuroscience*. 2002; 5(3):277–283. <https://doi.org/10.1038/nn816>
38. Hamann S. Cognitive and neural mechanisms of emotional memory. *Trends in cognitive sciences*. 2001; 5(9):394–400. [https://doi.org/10.1016/S1364-6613\(00\)01707-1](https://doi.org/10.1016/S1364-6613(00)01707-1)
39. Droit-Volet S, Meck WH. How emotions colour our perception of time. *Trends in cognitive sciences*. 2007; 11(12):504–513. <https://doi.org/10.1016/j.tics.2007.09.008>
40. Tielman M, Neerincx M, Meyer JJ, Looije R. Adaptive emotional expression in robot-child interaction. In: 2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI). IEEE; 2014. p. 407–414.
41. Houstis O, Kiliaridis S. Gender and age differences in facial expressions. *The European Journal of Orthodontics*. 2009; 31(5):459–466. <https://doi.org/10.1093/ejo/cjp019>
42. Jack RE, Garrod OG, Yu H, Caldara R, Schyns PG. Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences*. 2012; 109(19):7241–7244. <https://doi.org/10.1073/pnas.1200155109>
43. Sokolov AA, Zeidman P, Erb M, Pollick FE, Fallgatter AJ, Ryvlin P, et al. Brain circuits signaling the absence of emotion in body language. *Proceedings of the National Academy of Sciences*. 2020; 117(34):20868–20873. <https://doi.org/10.1073/pnas.2007141117> PMID: 32764147