Research Article

Esthetic interaction model of robot with human to develop social affinity

Jae-Joon Lee[1], Dae-Won Kim[1] and Bo-Yeong Kang[2]

Abstract

Esthetic interaction of robot with human is a new human–robot interaction model that focuses on helping a human create a social affinity with a robot by adding the concept of an esthetic experience to their interaction. The model is based on the theory that an esthetic experience is a circular mental process between subjects and their social environment; this experience makes social interaction more positive in ways that differ from more traditional emotion-related approaches. Our research using esthetic interaction of robot with human demonstrates that by including an esthetic interaction (imitative play) with the robot, the robot's negative emotional factors (e.g. facial expression and appearance) do not produce a negative social relationship with a human but transform the relationship into a positive one. The results do not signify that emotional factors can simply be excluded from human–robot interaction; rather, they suggest that there are circular processes between the emotional factors and the esthetic interaction with the robot. Although our research is challenging and very experimental, we expect it to contribute to innovation in human–robot interaction research fields.

Keywords

Human–robot interaction, sociable robot, esthetic experience, esthetic interaction

Introduction

Technological issues regarding the relationships between humans and robots continue to accumulate. Now, many people claim that the issues are humanistic as well as technological. Post-humanists S Zizekand B Stiegler said that, in the 18th century, modern humans could define themselves as rational beings by their own autonomy; now, however, we can define ourselves only by our relevance to technology.[1–3] If this is the case, human dependence on technology means that machines, as the realization of technology, are no longer simple tools, instruments for human use, or mere extensions of the human body. In technology-dependent environments, human intelligence, emotion, and sociability cannot be explained by humans themselves but only with the aid of technological artifacts: for example, our thought by Gutenberg's printing press, our sight by Daguerreotype cameras, our intelligence by artificial intelligence, and our body by functional magnetic resonance imaging.[4,5] Thus, interactions between humans and robots may be no exception.

Ecological approaches create a constructive and structural coupling between human and machine through their multiple, various interactions.[6,7] According to this, the interactivity underlying social relationships is contradictory: The interacting entities want to maintain their

[1] School of Computer Science and Engineering, Chung-Ang University, Dongjak-gu, Seoul, Republic of Korea
[2] School of Mechanical Engineering, Kyungpook National University, Pukgu, Daegu, Republic of Korea

Corresponding author:
Bo-Yeong Kang, Kyungpook National University, Pukgu, Daegu 702701, Republic of Korea.
Email: kby09@knu.ac.kr

independence from each other (including human and non-human); however, they simultaneously seek for connections with others to make up for their own ontological deficiencies. Therefore, social entities have their own experience but are related with others.

In fact, many unique meanings of life may be comprised of the organization of these socioecological experiences. Traditional esthetics has also examined the meaning of life using the concept of esthetic experience. Recently, researchers have claimed to have meaningful experiences with techniques in everyday life.[8,9] Therefore, we believe that some esthetic interactions between human and machine (intelligent robot) can give us a meaningful experience in socioecological conditions. The esthetic interaction in the present study is based on this type of experience.

Esthetic interaction can be differentiated from other interactions because it is always accompanied by feelings. An esthetic interaction does not produce an experience that leads to intellectual expansion, for example, in the aggregation, storage, or use of information or knowledge. Rather, it allows people to learn esthetic communicative forms from their own cultural communities and helps them find a refined communication for better social affinity. Thus, if the concept of experience by esthetic interaction is applied to human–robot interaction (HRI), we can expect a robot to develop more esthetically refined interactions and become more sociably advanced.

Because esthetic interactions generally include emotional states, most esthetic approaches to HRI are carried out like emotional approaches, which have been evaluated as a so-called "smoother" way to advance a robot's communicative abilities.[10,11] Breazeal and Hanson claim that positive emotions can help improve a robot's sociability when interacting with a human. Despite the emotions' importance, social affinity is not merely decided by positive emotions, especially in a practical communication situation. People often experience negative emotions (or emotionally neutral states of mind) as well as positive ones, in everyday life situations. Thus, if we conclude that positive emotions are the only key for understanding sociability, many social interactions may be overlooked. However, we know that even in a negative emotional condition, a person can successfully interact with others and feel positive emotions at the end of the interaction. Thus, the following questions attracted us. How can a negative interaction be transformed into a positive one? How can the transformation draw wonderful social interactions?

Esthetics suggests that this negative-into-positive communication mechanism can be explained with esthetic interactions, for example, mimicry play or imitative play ("mimesis" in a broad esthetic sense). In western societies, imitative play is generally used as a form of social communication based on feelings. While doing imitative play using gestures or facial expressions, partners repeatedly affect one another through their emotional esthetic experience and change their social attitude toward one another to suitably match their situation. In other words, the esthetic experience of playful imitative interactions can change a cognitive attitude from negative into positive, and thus can help one cope with a socially negative environment. Therefore, we can reasonably expect that esthetic experience processes can promote a robot's sociability in interactive situations with humans.

In a previous study,[12] we presented the first social-robot study to investigate child–robot interactions based on esthetics. The imitative interactions of the child and robot were based on Meltzoff–Moore's esthetic interaction model. A robot prototype was implemented and tested with 10 children for verification.

The main goal of this article is to present a theoretical esthetic model (esthetic interaction of robot with human (AIR-H)) that shows the schematic structures and technical platform of esthetic interaction. To the best of our knowledge, this is the first study to investigate a theoretical model for robot interaction based on esthetic theory. Our work in this article focuses on three issues that were not addressed in earlier work.

First, although the concept of esthetic interaction was used for HRI by Lee et al.,[12] no theoretical or concrete model was given on why esthetic judgment was more influential than emotions in playful interactions between the children and the robot. Thus, we clarify the concept of esthetic experience that we intend to apply to HRI. Then, we define the elements of an esthetic experience and explain the circularity of its processes.

Second, the prototypical robot has been upgraded with enhanced multimodal interactions, for example, more natural facial expressions, emotional gestures, and changes in the interaction scenarios. Third, through comprehensive experiments and analyses, we demonstrate the effectiveness of our approach. We performed on-site experiments with the robot and 37 children in a kindergarten. This task was difficult because several people were involved in the test to systematically conduct the experiments: children, their parents and teachers, and our staff.

The remainder of this article is organized as follows. In "Theoretical background of the esthetic interaction between human and robot" section, we explain the theoretical background for esthetic interactions between humans and robots. Specifically, "Esthetic experience and interaction" section clarifies the esthetic experience concept that we intend to apply to HRI. In "Generality, basic elements, and constructive processes of esthetic interactions" section, we define the simplified elements of an esthetic experience and explain the circularity of its processes. We show that the circularity corresponds to interactivity in the relationships between humans and robots.

"AIR-H schema, platform, and design" section presents the schematic structures of the esthetic interaction and the technical platform. We schematize the conceptual model for esthetic interactions between humans and robots in "Schematization of the AIR-H esthetic experience" section

and call the model "AIR-H." Our research suggests a new point of view, which differs from existing studies on emotional robots by applying esthetic concepts to interactive situations between a human and a robot.

The technical elements of the model's esthetic interactions are described in "Robot platform for AIR-H," "Facial and gestural expressions," "Facial expression recognition," and "Interaction scenarios" sections. The model addresses the limitations of other approaches that used positive emotional responses as the primary sociability factor in HRI. Our model alternatively proposes that esthetic experience through "imitative play" can effectively advance the sociability of the interaction between a robot and a human. In this study, we focus on the sociability of children by reducing the human psychological tension and inducing active behaviors with a robot. In "Experimental results" section, we discuss our experiment including a robot playing with children (mimicry play or imitative interaction) to validate the model and the technical design. "Conclusions and future work" section concludes the article.

Theoretical background of the esthetic interaction between human and robot

Esthetic experience and interaction

Generally, robot esthetics has interpreted the cultural meaning of robots philosophically; thus, it is uncommon to apply esthetic concepts to robots' technological issues. Nevertheless, a series of esthetic approaches exist in the robot interaction research field. The initial studies focused on the esthetic appearance of a robot.[11] The interest in a robot's esthetic appearance originated from Masahiro Mori's hypothetical doctrine of "the uncanny valley" and has influenced researchers both directly and indirectly.[13]

Good looks or a beautiful appearance are positive elements for successful social communications or interactions, and esthetic approaches consider them a common issue. However, esthetics treats appearances and an object's external form as ostensible or partial esthetic properties because those properties are only the starting points of the esthetic experience.

For example, people do not always respond the same way to the appearance of the same object, for example, "Yesterday, her beautiful dress caught my eye, but today somehow it looks dowdy. I guess it is more suited to her sunny personality. I think she looks more beautiful when she smiles." In this example, the subject's esthetic judgment connects the sensible information of the girl's appearance and his feelings about it with his sociopsychological attitude. People continually note the subtle differences in their sensible information, modify them, and obtain an esthetic experience from them. The construction of an esthetic experience is the result of long interactive processes with these parts.

According to Deweyan esthetics, an esthetic experience is "an experience" that people obtain meaningfully when they meet people, artifacts, or natural things, with feelings as the second-order consciousness of emotions or sensations.[14] It spiritually reinforces their lives.[8,9,15] The key advantage of the Deweyan concept is that esthetics is no longer restricted to works of art but extends to other topics in everyday life, such as robots. With the concept of an esthetic state of mind, estheticians can more clearly explain the Deweyan concept of an esthetic experience. It is defined as the inner (or phenomenological) state of a cognitive and emotional mind, which is constructed with esthetic objects.[16]

Meanwhile, an individual's esthetic experience is the esthetic state of mind composed of interrelations (or interactions) with others, that is, his/her environment, in the ecological sense. If someone's interactions with his/her ecological others provide him/her with a meaningful experience of esthetic conditions, the interactions can be termed an esthetic interaction. Therefore, an esthetic experience is redefined as an interactive state of mind that is constructed from meaningful feelings when an individual interacts with his/her others in an esthetic environment (e.g. play). These feelings become the basis for sociality through these interactions. Typically, esthetics considers that esthetic judgment is a onetime complete cognitive decision regarding an object. However, such determinism can change throughout a continuous (inter) relationship with the object. Esthetic experience involves a variable and constitutive state of mind. As such, the experience can be regarded as a changeable and interactive state of mind.

If a robot satisfies these conceptual conditions at a basic level, we can expect that an esthetic experience will be constructed through interactions between a human and the robot. The robot, as an esthetic object in Deweyan esthetics, can be expected to communicate socially with its human partner through interactions.

Generality, basic elements, and constructive processes of esthetic interactions

Generality. Esthetics tries to explain that subjective and individual feelings can be communicated generally among social members. Feelings are an inner state of mind that cannot be easily defined. However, people generally believe that subjective feelings can be shared. They once were explained as a relationship with linguistic symbols[17,18] or physiological signs.[19] Thus, the process of sharing a subjective state of mind is one of the crucial problems that esthetics must resolve with the concept of esthetic experience.

A solution for this problem is given when community members generally agree on feelings. Here, general agreement means that there are beliefs, emotions, attitudes, and tastes shared with social members who are in similar linguistic conditions.[20] Most social members think that they have different states of mind than others. Simultaneously, they believe that they are social beings with similar feelings
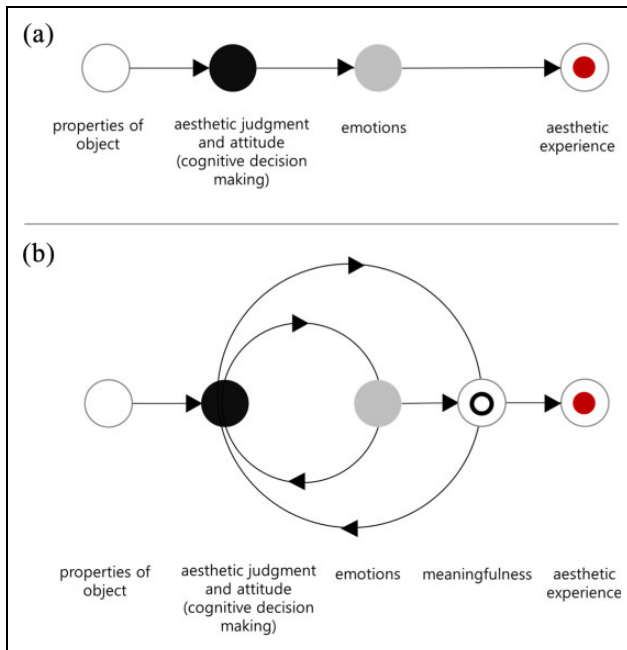
Figure 1. Linear and circular models of esthetic experience. (a) Traditional esthetics follows a linear model. (b) We propose a circular model and evaluate its efficiency for esthetic interactions with a robot.

through interactions in similar environments. This is called a socially felt18 belief, that is, a common sense. Such a belief is also based on the similarity of the human biological structure. The same species has similar biological structures and, therefore, has similar cognitive structures.[21]

To reduce the incongruity between the individual and his/her social community, or between the inner world and the outer world,[22] people continually try to revise their own feelings into socially agreeable forms. Therefore, feelings can be generalized using the social training of repetitively tuning one's feelings with the esthetic. The esthetic version of a socially felt belief is expressed as the esthetic, for example, beautiful and ugly, graceful and sublime, and so on. Eventually, the esthetic interaction becomes a continual training or learning process to acquire the social agreement of feelings and finally to reach an esthetic experience.

Basic elements. Esthetics typically defines the basic elements of an esthetic experience as follows: an object's esthetic properties; cognitive decision-making, which includes the perception of an object; emotional states that follow those decisions; and the meanings of the cognitive and emotional states. The relationships between these elements are determined by a simple linear causality model, as Kant's theory explains.[23] Namely, an object has some properties; an individual perceives and recognizes these properties, and then appraises whether the properties are in his/her favor. Based on the appraisal, the individual feels pleasure or displeasure; finally, this linear process causes an esthetic experience (Figure 1(a)).

Constructive processes. Empiricist D Hume considered an esthetic experience as the construction of a socially transferable mind and explained a prototypical esthetic experience concept, for example, taste, as opinions or feelings of unprejudiced people with good sense.[20,24] Hume said that taste was constructed through the "test of time,"[24] and this continual training meant repetitive interactions between esthetic subjects and their objects in social environments.

As mentioned earlier, the basic elements of an esthetic experience consist of object properties, a cognitive decision, and emotions. Someone perceives the object's properties, for example, its appearance, judges whether the newly perceived contents are appropriate to his/her existing images, and decides whether the new ones satisfy him/her. The emotional responses are then added to the decision. Up to this point, they are equivalent to the typical elements of an esthetic experience in the linear process model.

Now, however, we define a new esthetic experience, to be constructed using circular relationships between the basic elements. Because the influence of the emotional states on preceding cognitive states is restricted in the linear process, an object's esthetic experience ceases when it arrives at an emotional state after making a cognitive decision. In contrast, decisions can be changed by the emotional states, and their meanings are revised repeatedly in the process. Eventually, the esthetic process is recursively and repeatedly modified until the experience reaches a homeostatic state and includes numerous meanings. In the circular process, esthetic interactions are socially communicative actions that use feelings to reach this homeostatic state, and then generate socially transferable forms of the feelings. Therefore, an esthetic experience is redefined as the constructive state of mind that includes all of these recursive processes between the cognitive and emotional elements (Figure 1(b)).

Regarding these, appraisal theorist N Frijda insists that there is a reciprocal relationship between cognitions and emotions (Figure 2(b)). He considers the relationships as a purely inner process[25,26]; however, the relationship is connected with the outer world by the subjects' actions. According to R Lazarus, emotional appraisal is the circular cognitive process of adapting to social environments.[27–30] His ecological model of the cognition–emotion–adaptation system includes cognitive elements, for example, motivation, attribution, coping, and self-monitoring. These elements operate circularly when an organism interacts with its own objects or events in a given situation (Figure 2(c)).

Ecological psychologist J Piaget also insists that human intelligence in early childhood is developed primarily by the successful results of the child's social and ecological adaptation to his/her environment. Interactions are necessarily required for that adaptation to occur.[31] Therefore, an esthetic experience is not simply a cognitive decision regarding beauty, with the accompanying emotions. Rather, it can be said that an esthetic experience is the
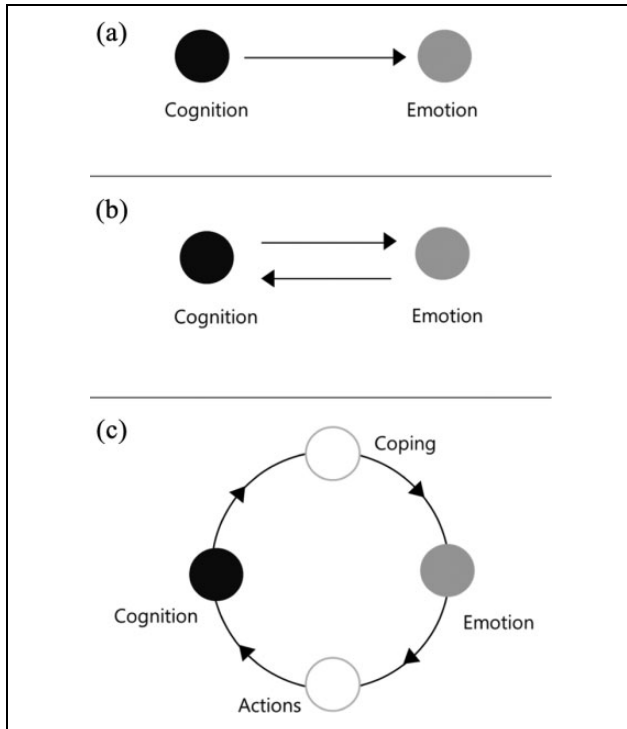
Figure 2. The circular model of the relationship between cognition and emotion effectively explains the adaptive interactions of social members in their social communication situations. We apply this circular model to our model of esthetic interaction between human and robot: (a) linear model of cognition and emotion, (b) reciprocal model of cognition and emotion, and (c) circular model of cognition and emotion.

Figure 3. The proposed schema of the esthetic experience process for the AIR-H model is based on circular, repetitive relationships between cognitive elements and emotions in a cognition–emotion–adaptation theory. AIR-H: esthetic interaction of robot with human.

result of a socioecological adaptation based on feelings, allowing an individual to effectively interact with others.

This long and recursive adaptive process of interactions supports the esthetic interaction as a kind of "play for play." This lasts for a relatively long term, and thus promotes a social affinity. "Play" (old English "pleġian") is an activity for enjoyment rather than a serious purpose and has a similar etymological origin as the theatrical performance as a genre of artistic form. Mimicking play or imitative play can be regarded as a prototypical form of esthetic interaction. Clearly, esthetics has used theatrical play as a social communication strategy for the satisfaction of society's cultural and educational interests.

Cognitive scientists A Meltzoff and S Baron-Cohen claim that infants read their caregiver's mind and embody the caregiver's cognitive and emotional mechanisms by imitating his/her facial expressions or gestures. In the process of playful imitation, although infants have no intention of learning social-communication skills, they effectively become sociable; even though they imitate their relational partners simply for pleasure itself.[32,33]

Based on this theoretical background, we can assign mimicry play a central role in the esthetic interaction between the children and a robot. It allows the children to
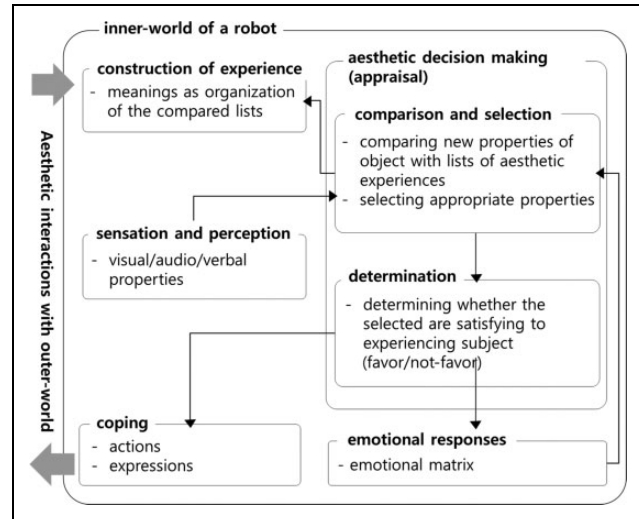
have esthetic satisfaction with positive emotions and presents a potential advancement in sociability between humans and robots.

AIR-H schema, platform, and design

Schematization of the AIR-H esthetic experience

The esthetic interaction between human and robot induces an intersubjective experience for the human based on feelings. If the interactions are in an esthetic form, like play, and simply cause pleasure without any intention or purpose, they can be referred to as esthetic. Finally, the contents of those esthetic interactions can be considered an esthetic experience. As we cannot determine whether a robot really has esthetic experiences, we consider the robot's esthetic experience from the experimenters' point of view. However, we expect that these esthetic interactions will improve the sociability between human and robot.

The proposed AIR-H model is based on the aforementioned theoretical foundations. Figure 3 shows the AIR-H conceptual schema. It includes the basic elements of an esthetic experience: object properties, cognitive decision-making regarding an object, emotional states presented after/before the decisions (circularly), and the meanings of the cognitive and emotional states. Actual interactions realize these elements following the circular procedure, and AIR-H is based on the recursive relationships between the cognitive and emotional aspects.

We verified the effectiveness of our AIR-H model through experimentation that studied the esthetic interactions between a child and a robot. In our experiment, the
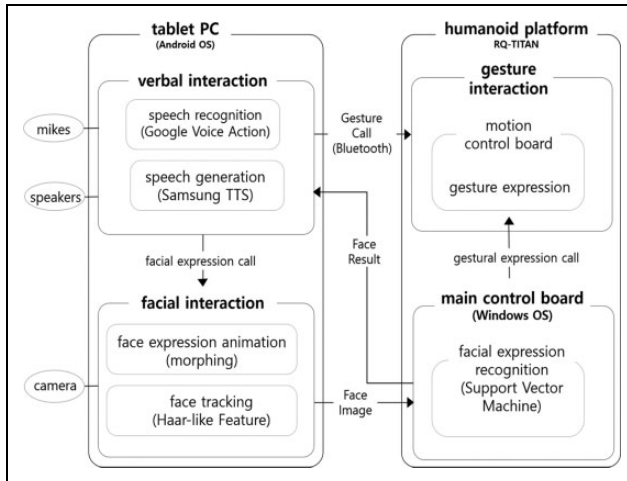
Figure 4. AIR-H robot implementation with multimodal interactions. AIR-H: esthetic interaction of robot with human.

terms were defined as follows: mimicry or imitative play is the esthetic interaction; the object's esthetic properties are the interaction partners' appearances (as well as verbal, gestural, and facial expressions); esthetic judgment (i.e. esthetic cognitive decision-making) is described as the favor/no-favor(lack thereof) in each partner; and the meanings of the cognitive and emotional responses are a type of social attitude toward each partner (e.g. "I like you. I want to see you again.").

In our model, we implemented gestures and appearance characteristics, for example, sweet, smiling, ugly, or angry faces, on a robot. Then, we allowed children to enjoy mimicry play with the robot and observed how the sociability between the children and the robot changed dynamically as they interacted. Mimicry play is a play for pleasure itself, without any premised social purpose. However, it causes positive emotions and improves a child's cognitive decisions regarding a social partner; here, a robot.

If a child, even one that meets the robot with a negative appearance, can be encouraged to participate in mimicry play with the robot, an esthetic experience can be constructed for the child intersubjectively, and the social relationship between the child and the robot will improve. Thus, we show that the traditional external properties of a robot, such as its appearance, are not crucial factors in the construction of HRI sociability and prove that esthetic interactions through mimicry play are effective in constructing sociable relationships between humans and robots.

Robot platform for AIR-H

Based on the AIR-H conceptual schema shown in Figure 3, the AIR-H robot platform was implemented as shown in Figure 4. The AIR-H robot hardware is mainly composed of a humanoid RQ-TITAN[34] for the body, and a tablet PC, the Samsung Galaxy Tab 7.7, for the face. The robot
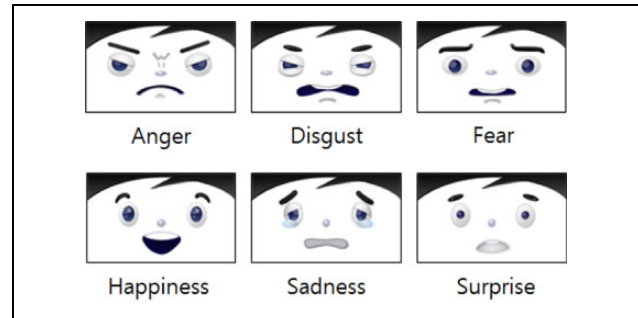

Figure 5. Facial expressions designed using Ekman's six emotional states.

software is composed of three interaction components: facial, gestural, and verbal. It also effectively supports the child's emotional representations using multimodal interactions.

The facial interaction component in AIR-H consists of one subcomponent that displays facial expressions for imitative interaction and another subcomponent that recognizes whether the child imitates the facial expressions. Using the facial interaction component, the robot generates the facial expressions as animations on the tablet PC. To identify whether the child imitates the robot's facial expression, the robot first tracks the child's face, then recognizes the child's facial expression. The result of recognizing the child's facial expression is delivered to the verbal and gestural interaction components, as well as to the facial interaction component, so that the robot shows an appropriate response to the child's imitative interaction result.

The verbal interaction component allows the robot to converse with the child and is divided into speech recognition and speech generation modules. We used the Google voice action APIs for speech recognition on the child's speech input through a microphone, and the Samsung text-to-speech engine for the robot's speech generation.

The robot's gestures increase the interaction quality by adding a body-kinesthetic factor in addition to the visual and auditory factors, thus enhancing the child's esthetic experience. The gestural interaction component calls ready-made gestures, saved in a control board, when the verbal interaction needs a gesture during the conversation.

Facial and gestural expressions

Face design with emotional expressions. Facial expressions are central to the study of emotional states. Of the various theories on human emotion, we used Ekman's six emotional states[35,36] for our facial expression design: anger, disgust, fear, happiness, sadness, and surprise. The robot's facial expression was implemented as an animation on a tablet PC; each facial feature, for example, eyes, nose, lips, and eyebrows, was designed separately in each emotional state, so the robot could have various facial expressions.

Table 1. Gestural expressions in imitative interaction.

Situation	Emotional states	Gesture
Welcoming or greeting	PG	Swing an arm to the right and left lightly and slowly
Unfriendly reception of a guest	NG	Half-raise and swing an arm strongly and quickly
Playing with a child	PR	Swing arms forward or backward rhythmically to music
Pointing out a mistake	NR	Swing arms to the right and left strongly and quickly
Suggesting	PP	Stretch and half-raise arms forward slowly

PG: positive greeting; NG: negative greeting; PR: positive reaction; NR: negative reaction; PP: positive proposition.

Figure 5 shows the designed facial expressions for the six emotional states. The happiness expression corresponds to a positive emotion, while the other expressions correspond to negative emotions.

Seamless transitions between facial expressions were provided using morphing, a special effect in animation that changes one image to another through interpolation. In this case, we used FantaMorph,[37] a well-known commercial application, to offer professional quality animation.

Gesture design with emotional expressions. Since gestural expressions, like facial expressions, are closely linked to emotional states, gestures have been studied in various emotion-related contexts.[38–40] We considered 12 types of general body movements. Stretching, opening, upward, forward, light, and slow movements can elicit positive emotions, such as interest and joy, while bowing, closing, downward, backward, strong, and fast movements can elicit negative emotions, such as anger, contempt, and disgust.

We composed our gestural expressions based on these body movements and their implied emotions. The combination of body movements is shown in Table 1; they represent the emotional states in the imitative interaction situation. For example, when the robot meets the child and greets him in a friendly manner, it stretches its arm in an upward and open posture and swings it lightly and slowly. On the other hand, to express a negative feeling and an unfriendly greeting, the robot stretches its arm to a relatively lower position with a closed posture and swings it strongly and quickly.

Facial expression recognition

In our research, the aim of the facial expression recognition component is not to measure how exact the child's imitation is but to increase the communication reliability by informing the child that the robot can identify the child's facial expression. After recognizing the child's facial expression, the robot can show the appropriate response to compliment or encourage the child based on the recognition results.

The facial expression recognition component was implemented using a support vector machine,[41] which can recognize whether the child imitated the positive facial expression as positive and the negative one as negative. The system tracks the child's face (100 × 100 pixels) using the Haar-like feature[42] from an image input through a camera on the tablet PC. After extracting the lip feature from the lower region (60 × 30 pixels) of the face image, the system classifies the face's emotion as positive or negative. As learning data, we collected 200 positive and 200 negative face images from the Web. The recognition system showed 91.25% classification accuracy for these 400 images after 10-fold cross-validation.

Interaction scenarios

In accordance with our assumption that an esthetic interaction has a more crucial impact on the improvement of social interactions than a simple emotional interaction, the following imitative scenarios were developed to compare a child's sociability in an emotional interaction with that of one in an esthetic interaction.

A. Simple imitative interaction scenario
- Positive scenario (Pos): The robot exhibits positive facial, verbal, and gestural expressions to evoke a positive emotion.
- Negative scenario (Neg): The robot exhibits negative facial, verbal, and gestural expressions to evoke a negative emotion.
B. Playful esthetic imitative interaction scenario (AES)

- The robot exhibits both positive and negative facial, verbal, and gestural expressions continually and rhythmically.

We designed a simple imitative interaction scenarios (Pos and Neg) to test the previous approaches on the emotional interaction, in which positive emotional expressions can be thought to enhance positive social attitudes, while negative emotional expressions can be thought to enhance negative social attitudes. Hence, in the Pos, positive facial expressions, speech, and gestures were presented, and in the Neg, negative ones were presented.

The negative simple imitative interaction scenario is presented in Figure 6. In this scenario, the child imitates the robot's negative facial expression, and the robot responds with slightly negative emotional expressions, even if the child imitates well; if the child fails to imitate, it shows a strong negative response to require the child to attempt the imitation again. The child can repeat this procedure as often as needed.

Step	Speech and actions	Face	Gesture
1:Child	*The child comes into the laboratory with his/her mother or teacher.*		
2:Robot	*The robot wakes up and opens its eyes.* Hello. My name is Chulsu. What is your name?	Neutral	NG
3:Child	I'm 000.		
4:Robot	Why did you come here?		
5:Child	000		
6:Robot	Hey, let's imitate my face. (Alternatives: And how about this? Can you imitate me again? Etc.) *The robot suggests a facial expression.*	 N	
7:Child	*Child imitates the robot's face.* (If the child imitates well, goto step 8, otherwise, goto step 8-1.)		
8:Robot 8-1	Hum, you did it (negatively). (Goto step 6.) No, no! Try again. (Goto step 6.)	N N	NR NR
	I don't want to play with you anymore. Goodbye.		 NG
9:Child	Bye, see you again.		

Figure 6. Simple imitative interaction scenario (negative scenario).

A playful AES exhibits the robot's positive and negative emotional expressions to a child. In this scenario, the interaction was changed from a simple imitation with only emotional interaction to a play pattern imitation with both esthetic and emotional interactions. Figure 7 shows the playful AES and Figure 8 shows a sample combination of the positive and negative facial expressions used for an esthetic imitation. In step 4-2 of the esthetic scenario shown in Figure 7, three facial expressions that combine positive and negative emotional states are rhythmically presented to the child as a play pattern, and the child imitates them. After the repetitive imitation interactions, we ask for an esthetic appraisal from the child, using questions such as "Did you have fun?"; "Am I pretty?"; and "Do you like me?" (steps 6-1, 8-1, and 10-1, respectively, shown in Figure 7).

The children underwent one of two procedures: Pos followed by AES, or Neg followed by AES, during which we tried to identify whether the repetition of a playful esthetic imitation resulted in an esthetic experience between the child and the robot. Furthermore, we attempted to verify that this esthetic interaction

experience could change or enhance the child's social attitude of the child toward the robot. In all of the scenarios, the facial expression was recognized. The robot showed positive responses if the child imitated well and showed negative responses or provided suggestions for improvement if the child imitated poorly.

Experimental results

Experimental setup

To verify the effectiveness of the proposed AIR-H model in enhancing children's sociability, we selected 37 kindergarten children, ages 4–5 years, to participate in this study. The experiment was carried out in the kindergarten classroom to ensure that the children were familiar with and felt comfortable in the environment. Figure 9 shows a snapshot of the esthetic interaction experiment between a child and the robot. The child interacts with the robot across the table, and an observer sits beside the child to coordinate the test, including greeting the child before the actual experiment and questioning the child following their interaction with

Step	Speech and actions	Face	Gesture
1:Child	*The child comes into the laboratory with his/her mother or teacher.*		
2:Robot	*The robot wakes up and opens its eyes.* Hello. We met a little while ago, right? Do you know my name? Tell me my name.	P	PG
3:Child	*The child answers.* (If the child remembers the robot's name, go to step 4, otherwise goto step 4-1.)		
4:Robot 4-1 4-2	You are so smart! Oops, my name is Chulsu. It's ok. Let's stay close friends. Hey, do you want to play with me? Let's imitate my face. (Alternatives: And how about this? Can you imitate me again? Etc.) *The robot shows three emotionally dynamic facial expressions, one after the other.*	 P-P-N	 PP PR,NR
5:Child	*Child copies the robot's face.*		
6:Robot 6-1	Wow, you are doing very well! (Alternatives: Good job! Great! Etc.) (Goto step 4-2.) 000, did you have fun?		
7:Child	Yes, I did (or No, I didn't).		
8:Robot 8-1	I had fun! Am I pretty? Could you tell me?		
9:Child	Yes, you are pretty (or No, you are ugly).		
10:Robot 10-1	Ha-ha. Yes, I see. Then, do you like me?	P	
11:Child	Yes (or No).		
12:Robot	Yes, I see. I like you very much. By the way, do you remember my name? Could you tell me?		
13:Child	*The child answers.* (If he remembers the robot's name, goto step 14, otherwise goto step 14-1.)		
14:Robot 14-1	You are my best friend. I love you! That's ok. My name is Chulsu. Friend! It's time to say goodbye. See you later.	P	 PG
15:Child	Bye, see you again.		

Figure 7. Playful AES. AES: esthetic imitative interaction scenario.
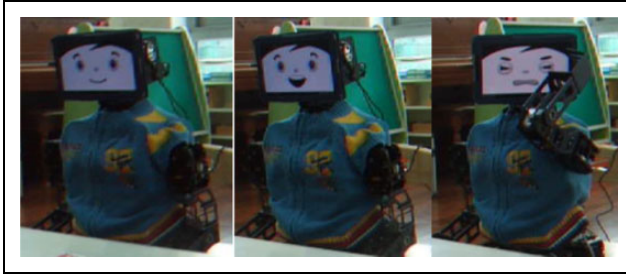
Figure 8. Sample esthetic imitative interaction sequence: smile–happiness–disgust.
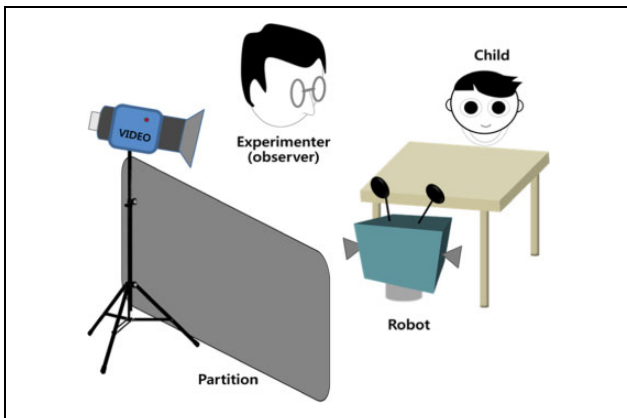

Figure 9. Esthetic interaction between a child and a robot (right). Experimental room setup (left).

the robot. A hidden video camera was installed, so that the child did not realize that they were being video recorded.

Table 2 lists the demographic and experimental characteristics of the children. There were 25 male and 12 female children, with a mean age of 5.1 years. To examine whether the children's social attitude (affinity) changed through the playful imitative interaction, we divided the 37 children into two groups. In the first group, 19 children participated in the Pos and Pos + AES experiments, in which they first performed simple positive imitative interactions with the robot (Pos) and then, after a short rest, underwent an esthetic experience (Pos + AES). In the second group, 18 children participated in the Neg and Neg + AES experiments, in which they first performed simple negative imitative interactions, and then experienced a playful esthetic interaction.

After conducting the four experiments (Pos, Pos + AES, Neg, and Neg + AES), we observed the facial expressions and gestures of the children to determine their social states; we evaluated three social states in the experiments: active, nervous, and imitative responses. Each child's social state was evaluated and scored by three raters, professionals in the field of HRI research.

For the initial evaluation, each professional gave a score on a five-point scale. For example, to determine how actively the child interacted with the robot, we gave a score

Table 2. Demographic characteristics of 37 children.

Characteristic	Category	Number of children (%)	Mean
Gender	Male	25 (68)	
	Female	12 (32)	
Age			5.1
Test type: Pos and Pos + AES		19 (51)	
Test type: Neg and Neg + AES		18 (49)	

Pos: positive scenario; AES: esthetic imitative interaction scenario; Neg: negative scenario.

of -2 (very inactive), -1 (inactive), 0 (neutral), 1 (active), or 2 (very active). The scores often varied from person to person; therefore, to derive a fair agreement between ratings, we adopted a majority voting process to reach a consensus. For example, a child's response was rated as inactive if two raters classified it as inactive. Through this negotiation process, the evaluation scores were rescaled to -1, 0, and 1. Here, 0 denotes a neutral response from the child; $+1$ denotes an active, nervous, or imitative response; and -1 denotes an inactive, not nervous, or nonimitative response, depending on the characteristic being judged.

Improvements in the social attitude of the children's responses

The subjects who participated in our experiments were children who were either socially active or socially passive. We compared the children's responses to the three social states (active, nervous, and imitative response). This allowed us to confirm differences in esthetic interaction according to subjects' social attitudes. Figure 10 shows the frequencies of the children's observed social states for each test. In Figure 10 (top), we see that most of the children who participated in the Pos and Pos + AES experiments had active responses, whereas only a small number of children in the Neg experiment gave active responses. However, the number of active children in the negative test increased after they experienced a playful interaction in the Neg + AES experiment. Figure 10 (middle) shows the number of children who gave nervous responses in each experiment. We observed that many children in the Pos and Pos + AES experiments were not nervous. In contrast, most of the children in the Neg experiment showed nervous responses. It is interesting to note, however, that the number of children who showed nervous responses was remarkably reduced in the Neg + AES experiment, even though negative facial expressions and gestures were presented to the children during the experiment.

This tendency can also be observed in the changes in the children's imitative responses, shown in Figure 10 (bottom). Most of the children imitated the facial expressions and gestures of the robot in the Pos and Pos + AES experiments, whereas the children in the Neg experiment did not show much imitative response. However, once again, we
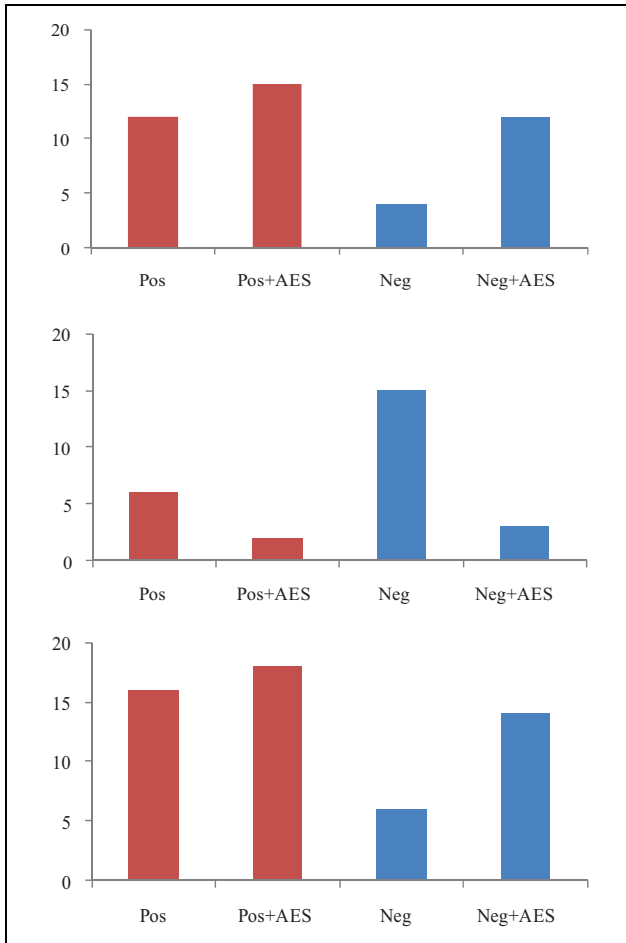
Figure 10. Frequency of a child's response: active response (top), nervous response (middle), and imitative response (bottom).

Figure 11. Number of children who remembered the robot's name successfully.
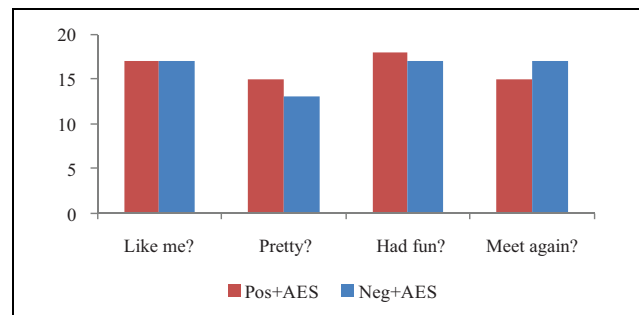
Figure 12. Number of children who answered Yes for the robot's questions.

observed that the number of children who displayed imitative behavior in the Neg experiment increased after the esthetic interactions in the Neg + AES experiment. Despite the previous negative interaction, they played with the robot by mimicking its facial expressions and gestures. The test results demonstrated that the children's social responses could be significantly influenced by esthetic interaction with the robot. Esthetic imitative interactions induce positive emotions and, accordingly, the positive emotions make children more interested, leading to the enhancement of their social attitudes.

To determine whether esthetic interaction enhanced the children's sociability, we also tested whether the children could remember the robot's name after interacting with it. Remembering another's name is a basic factor that is commonly used to identify the sociability level among members of a group. Figure 11 shows the number of children who remembered the robot's name. In the Pos and Neg experiments, the children did not often remember the robot's name. In contrast, many children correctly remembered the robot's name after playing with it in the Pos + AES and Neg + AES experiments.
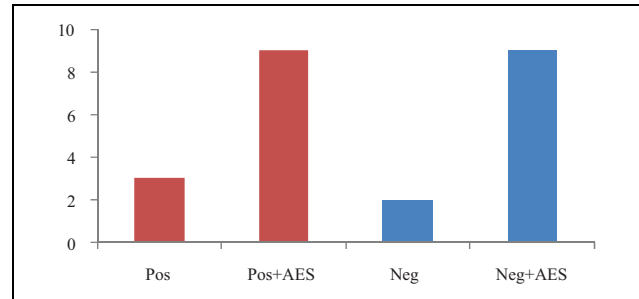
Generally, if the number of interactions increases in social behaviors, the opportunity to remember the social partner's name will increase proportionately. However, even though remembering the partner's name is not expected after only two interactions, our esthetic interaction experiment results showed that our subjects could recall the name of the robot. Moreover, the frequency of name recall for subjects participating in imitative play increased significantly in subjects with passive social personalities.

In addition, we asked the children four questions after the Pos + AES and Neg + AES experiments: (1) "Did you like the robot?"; (2) "Was the robot pretty?"; (3) "Did you have fun?"; and (4) "Do you want to meet the robot again?" The first three questions were also asked by the robot. As indicated in Figure 12, most of the children responded "Yes" to these questions after the esthetic experiments. All of the children who experienced a playful esthetic interaction with the robot stated that they had fun.

Tests of statistical significance

To make this study more informative, we assessed the differences in the children's responses using a statistical test. We used Fisher's exact test to verify if there were nonrandom associations between two response variables. Fisher's exact test is a well-known statistical significance test used in the analysis of contingency tables.[43] Using

Table 3. Contingency table of the children's active or inactive responses.

Response	Group			Group		
	Pos	Neg	Total	Neg	Neg + AES	Total
Active	12	4	16	4	12	16
Inactive	7	14	21	14	6	20
Total	19	18	37	18	18	36
Fisher test	*p* Value = 0.020			*p* Value = 0.018		

Pos: positive scenario; AES: esthetic imitative interaction scenario; Neg: negative scenario.

Table 4. Contingency table of the children's nervous or not nervous responses.

Response	Group			Group		
	Pos	Neg	Total	Neg	Neg + AES	Total
Nervous	6	15	21	15	3	18
Not nervous	13	3	16	3	15	18
Total	19	18	37	18	18	36
Fisher test	*p* Value = 0.003			*p* Value = 0.0		

Pos: positive scenario; AES: esthetic imitative interaction scenario; Neg: negative scenario.

Table 5. Contingency table of the children's imitative or nonimitative responses.

Response	Group			Group		
	Pos	Neg	Total	Neg	Neg + AES	Total
Imitate	16	6	22	6	14	20
Not imitate	3	12	15	12	4	16
Total	19	18	37	18	18	36
Fisher test	*p* Value = 0.003			*p* Value = 0.018		

Pos: positive scenario; AES: esthetic imitative interaction scenario; Neg: negative scenario.

this analysis, we can see if any differences in the observed social responses are significant. Tables 3 to 5 list the contingency tables of the children's responses in the different experiments.

In the left-hand side of the contingency table in Table 3, we observe that the Pos experiment has 7 inactive and 12 active responses, while the Neg experiment has 14 inactive and 4 active responses. The *p* value is 0.02, which is less than 0.05, and therefore, is significant. In this case, there is a statistically significant association between the response and the type of experiment. In the right-hand side of Table 3, we observe that the children's responses in the Neg experiment were quite different from those in the Neg + AES experiment, despite the negative facial expressions and gestures being presented to the children in both cases. The difference in responses is statistically significant because the *p* value is 0.018. We conclude that the esthetic experience helped evoke active responses in the Neg + AES experiment.

Table 4 shows the contingency table of the children's nervous and not nervous responses in the different experiments. Similar to the previous case, we see that the differences in the children's nervous responses between the Neg and Neg + AES experiments are significant (see the right-hand side of Table 4). The decrease in the nervous responses through the esthetic interaction is statistically meaningful (*p* value is 0.0) and allows us to conclude that the esthetic interaction plays a greater role than the emotional factors. This tendency is also observed in the analysis of the contingency table for the children's imitative responses (Table 5). We observe that the Neg experiment has six imitative responses and the Neg + AES experiment has 14 imitative responses, which is statistically significant (the *p* value is 0.018). This statistical test demonstrates that esthetic interactions are more influential than simple imitative interactions between a child and a robot.

Social affinity movement through esthetic interaction

We observed that playful esthetic interaction affected the children's social state. Hence, to examine the improvements in their social affinity or social attitude, we plotted their active–nervous (A-N) responses on a two-dimensional plane and compared their changes to facilitate visual inspection. Figure 13 shows the extent to which each response changed as a result of the playful esthetic interaction. The *x*-axis represents the degree of active responses as: active $(+1)$, neutral (0), and inactive $((-1)$. The *y*-axis represents the degree of nervous responses as: nervous $(+1)$, neutral (0), and not nervous (-1). Thus, the child's social state is represented as a point in A-N bipolar coordinates.

Multiple children can be located at the same (x, y) coordinate point if they gave the same response. Therefore, we use a filled circle to indicate the number of children stacked at a given point. A large circle indicates that many children showed this particular response. For instance, in the first chart of Figure 13, two children (in a small circle) showed inactive and nervous responses $(-1, +1)$, while 10 children (in a large circle) showed active and not nervous responses $(+1, -1)$.

The first and second charts in Figure 13 show the response changes of the 18 children who participated in the Pos and Pos + AES tests, respectively. We observe that the number of socially inactive children in the Pos test decreased with the esthetic interaction in the Pos + AES test. The Pos + AES test generally showed growth in the number of socially active children: 13 children gave active and not nervous responses (quadrant IV of the second chart). This tendency is more clearly observed in the Neg and Neg + AES tests, which are represented in the third and fourth charts in Figure 13, respectively. In the Neg test, six children, located at $(-1, +1)$, and seven children, located at $(0, +1)$, showed socially passive (inactive and nervous) responses because of the negative emotions shown by the robot. However, after
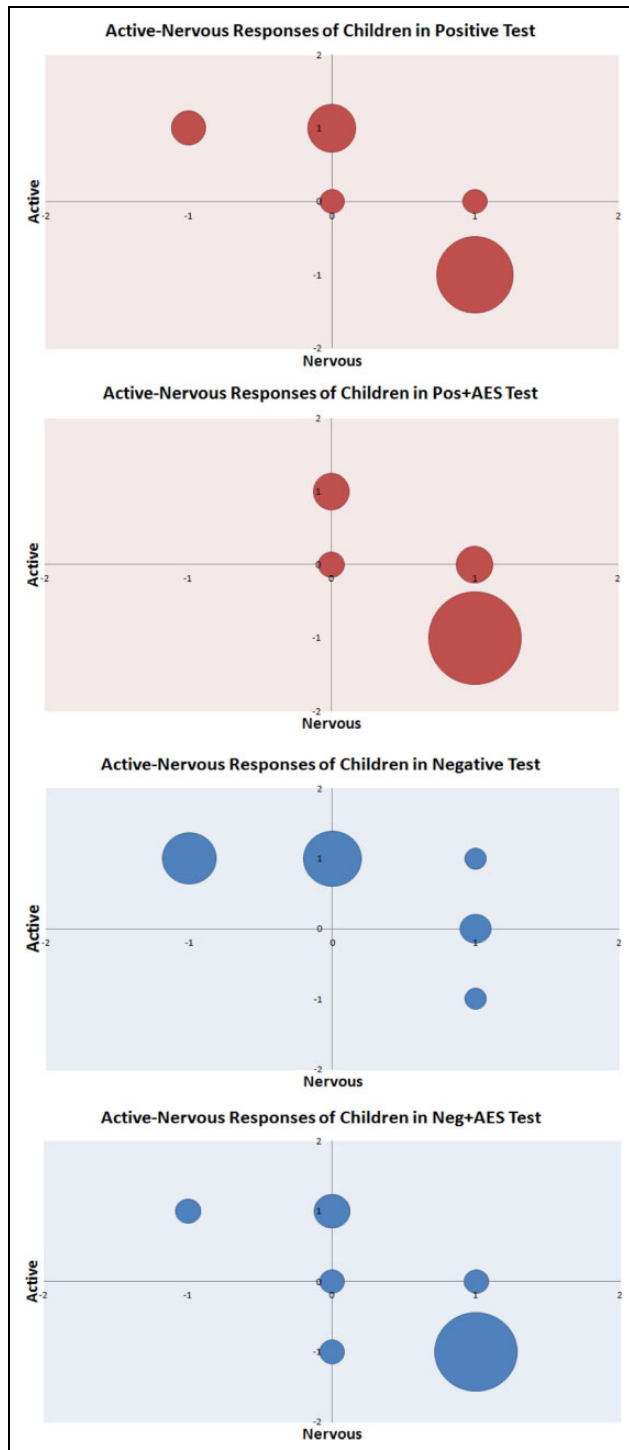
Figure 13. Changes in children's A-N responses. A-N: active–nervous.

experiencing a playful esthetic interaction, the responses dramatically improved toward quadrant IV (active, not nervous): 11 children showed active and not nervous social attitudes in the Neg + AES test.

To investigate the relationship between esthetic experiences and social affinity in a child–robot interaction, we conducted a survey of the children to find out their personality or character. One of the key questions we asked the parents about their children was whether their child tended to hide from new people they met. The other question we asked was whether their child tended to like robots. To clearly distinguish the results, we plotted the children's responses on a bipolar coordinate system using the five-point rating score of the observer who sat beside the children, since his proximity to the child allowed this individual to make more precise and fine-grained judgments.

Charts (a) and (b) in Figure 14 show the changes in the responses of the children who hide from strangers. Charts (c) and (d) show the responses of the children who do not hide from strangers. We can see in the Pos test, charts (a) and (c) in Figure 14, that the passive children (those who tended to hide) gave more nervous responses than those who did not hide. In contrast, in the Pos + AES test, charts (b) and (d), the social attitudes of the passive children were improved through esthetic play to be as active (+2) and not nervous (−2) as the children who did not hide.

Figure 15 shows the improvement of the A-N responses in the children who did not like robots. We see in charts (a) and (c) that the children who liked robots showed fewer nervous responses than those who did not. The children's responses to the Pos + AES test are noteworthy since the children who disliked robots showed interest in playing with the robot in chart (b). Their A-N responses improved, to be the same as those of the children who liked robots. Based on the experimental results, we concluded that the esthetic interaction between the children and the robot strengthened their positive social relationships and enhanced their social adaptability.

Conclusions and future work

When we made the children interact (playful and social interaction) with a robot rather than a human, we observed a social state in which the children thought of it as a friend, not a thing, and played with it naturally. Our research explained this phenomenon using the following two observations.

Firstly, emotional states based on emotional models are not ideal and unique conditions for a sociable robot but simply elements of its other conditions. However, many researchers in the field have considered positive emotional expressions as a universal condition for a robot's sociability. Our experiments revealed that negative emotional expressions in social relations between a human and a robot sometimes brought about positive influences in the relation that caused the esthetic experience. Our statistical significance tests, which used three degree measures (active, nervous, and imitative play), supported this. Esthetic interactions with the robot reduced the participants' nervousness and relatively enhanced their responses of playful imitation.
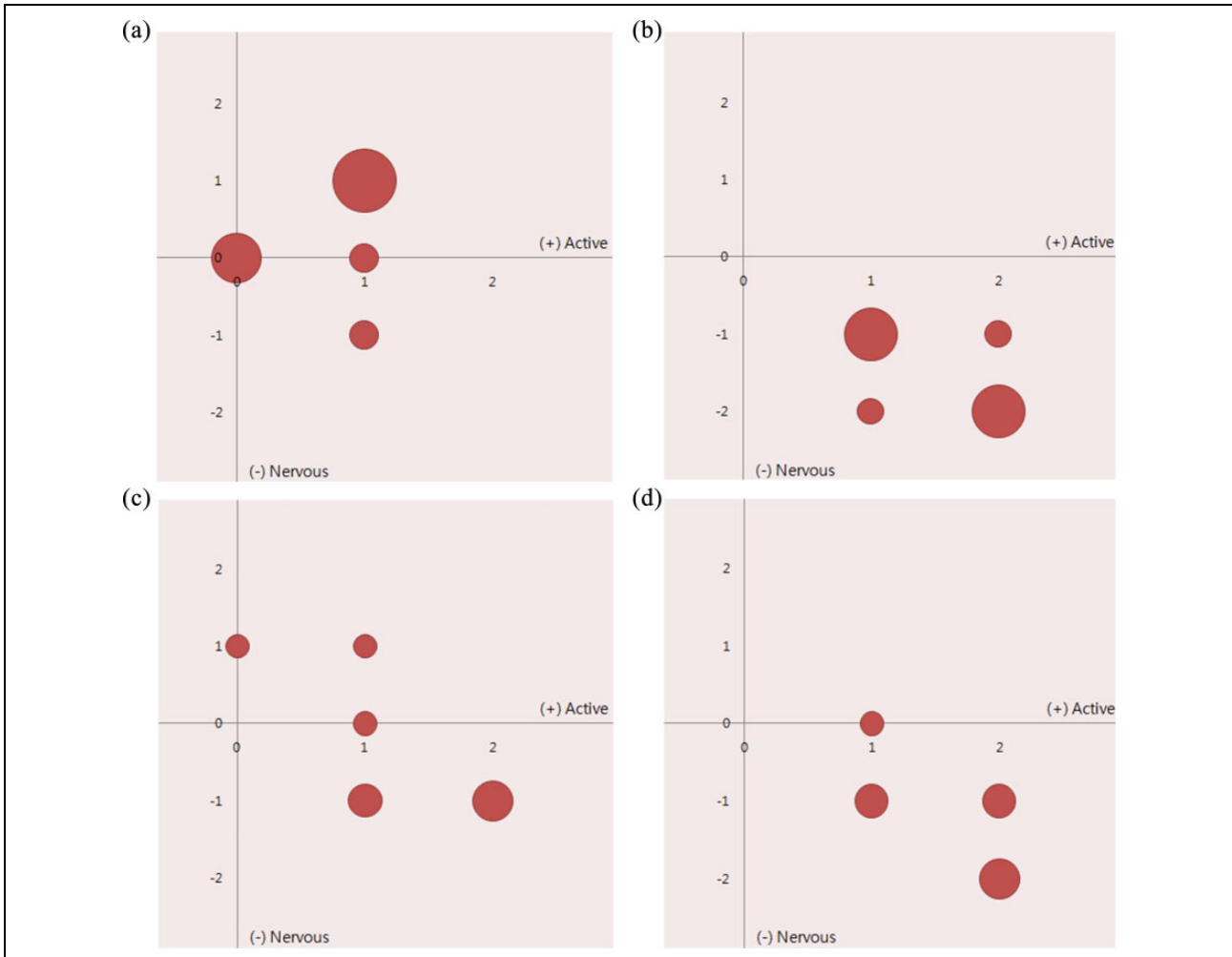
Figure 14. Changes in children's A-N responses. (a) Pos test with children who hide. (b) Pos + AES test with children who hide. (c) Pos test with children who do not hide. (d) Pos + AES test with children who do not hide. A-N: active–nervous; Pos: positive scenario; AES: esthetic imitative interaction scenario.

Second, it is clear that the emotional state caused by one's first impression of another will influence the social relationship. However, the esthetic concept of a robot's sociability is not restricted to its appearance. Our research transformed these very impressions into an esthetic experience that was constructed continuously through interaction with the robot. The experience was transformed through a recursive and repetitive interaction process between cognition and emotion (and also between the human and the robot).

The experiments showed that, although most participants showed a negative social attitude in the Neg experiment, their cognition from the esthetic experience under the Neg + AES conditions changed their social attitude to a positive one. They even exhibited positive emotional states when the robot displayed negative facial expressions. For a more sociable robot, our research explained a circular process model of esthetic interaction and implemented it in the interactions between a human and a robot. The implementation went beyond the common trend of emotion-based research regarding robots' sociability.

Although we could not determine whether the robot really had an esthetic experience, as mentioned above, we interpreted its inner state as if it had and observed that its sociability had advanced in the esthetic interaction conditions. However, the "observer's viewpoint" of a cybernetic conception is beyond the scope of this study and requires more complex theoretical elaboration in future work.

In our experiments, we interviewed many participants before the test to collect information on their personalities, and then closely examined whether their personalities (e.g. active, passive, shy, etc.) might influence their interaction with the robot. For example, an active personality might catalyze an individual to engage in social interactions more eagerly than others, while someone who felt psychological strain might avoid social interaction with strangers. We believe that these personality factors are very
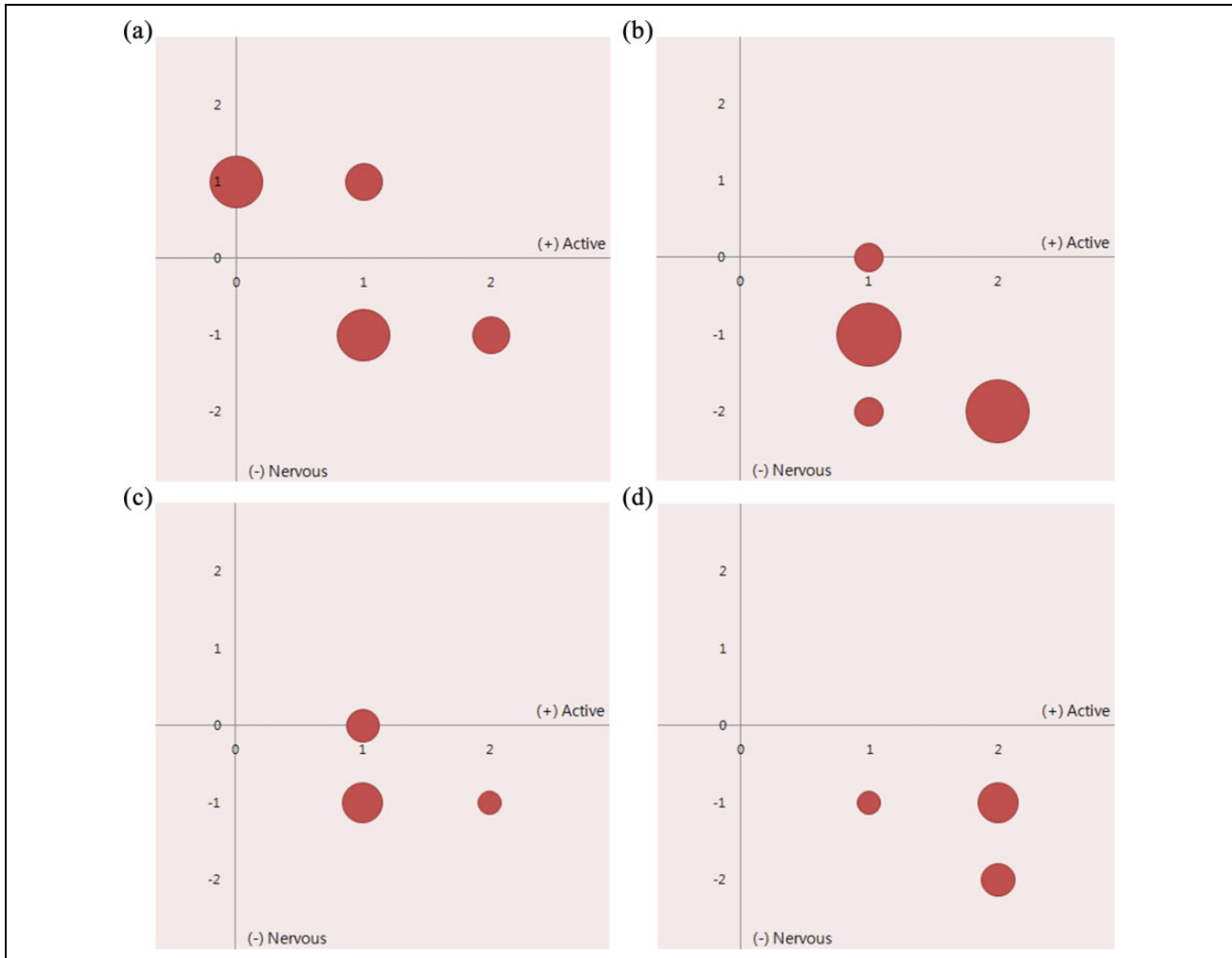
Figure 15. Changes in children's A-N responses. (a) Pos test with children who do not like robots. (b) Pos + AES test with children who do not like robots. (c) Pos test with children who like robots. (d) Pos + AES test with children who like robots. A-N: active–nervous; Pos: positive scenario; AES: esthetic imitative interaction scenario.

important in determining a human's social attitude toward a robot. This social attitude will suggest what social skills will be required to coexist with human beings for non-human (such as robots) in the near future society. We hope to explore them further in our future HRI research.

Our experiments focused primarily on children. Ultimately, however, our model of esthetic HRI may involve adults. Of course, there is a significant gap in the number and diversity of social behaviors for children and adults; however, in our future research, we intend to extend the application of the model to technologies for adults.

Lastly, the difficulties of studying interactions between humans and robots arise from the fact that the interactions are related to complex and delicate properties of human experience. Interaction studies in general will be able to explain these properties on many levels.

Nevertheless, if the levels can be largely divided, approaches exist for explaining the human aspects of interaction based on physiological signals in the human body.

Other approaches analyze the meanings of communicative interactions based on the social or emotional attitudes in conversations or behaviors.

Perhaps the multimodal approach in interaction will be related to the former. Of course, these two approaches can be interconnected, but we basically intend to analyze the interactions that are present in the latter. Toward that end, we aim to explain the qualitative dimension of interactions rather than the multimodal dimension (e.g. kinesthetic, tactile, visual, etc.) of bodily reactions, through interview analysis. Nonetheless, we understand the utility of analyzing physiological responses and will continue to do so in our future work.

Author note

Jae-Joon Lee is now affiliated to Sookmyung Research Institute of Humanities, Sookmyung Women's University, Yongsan-gu, Seoul, Republic of Korea.

References

1. Zizek S. *Organs without bodies: on Deleuze and consequences*. Routledge, 2012.

2. Stiegler B. *Symbolic misery, vol. 1: the hyper-industrial epoch*. Polity Press, 2014.

3. Latour B. *Science in action*. Harvard University Press, 1987.

4. Clark A and Chalmers DJ. The extended mind. In: Menary R (ed) *The extended mind*, MIT Press, 2010, pp. 27–42.

5. Clark A. *Being there: putting brain, body, and world together again*. MIT Press, 1998.

6. Pfeifer R and Bongard J. *How the body shapes the way we think: a new view of intelligence*. MIT Press, 2007.

7. Dourish P. *Where the action is: the foundations of embodied interaction*. MIT Press, 2004.

8. Dewey J. *Art as experience*. Southern Illinois University Press, 2008.

9. McCarthy J and Wright P. *Technology as experience*. MIT Press, 2004.

10. Breazeal C. *Designing sociable robot*. MIT Press, 2004.

11. Hanson D. Exploring the aesthetic range for humanoid robots. In: *Proceedings of the ICCS/CogSci-2006 long symposium: toward social mechanisms of Android science*, 26–29 July 2006.

12. Lee J, Kim D, and Kang B. Exploiting child-robot aesthetic interaction for a social robot. *Int J Adv Robot Syst* 2012; 9(81): 1–9.

13. Mori M. Bukimi no tani (the uncanny valley). *Energy* 1970; 7(4): 33–35.

14. Damasio A. *Looking for Spinoza, joy, sorrow, and the feeling brain*. Harcourt, 2003.

15. Alexander T. The art of life: Dewey's aesthetics. In: Hickmaned L (ed) *Reading Dewey: interpretations for a post-modern generation*. Indiana University Press, 1998, pp. 1–22.

16. Iseminger G. The aesthetic state of mind. In: Kieran M (ed) *Contemporary debates in aesthetics and the philosophy of art*. Blackwell, 2006, pp. 98–110.

17. Russell JA. Core affect and the psychological construction of emotion. *Psychol Rev* 2003; 110(1): 145–172.

18. Russell JA. A circumplex model of affect. *J Personal Soc Psychol* 1980; 39(6): 1161–1178.

19. LeDoux J. *The emotional brain: the mysterious underpinnings of emotional life*. New York: A Touchstone Book, 1996.

20. Hume D. Of the standard of taste. In: Miller EF (ed) *Essays, moral, political, and literary*. Cosimo Inc., 2006, pp. 231–255.

21. Varela F, Rosch E, and Thompson E. *The embodied mind: cognitive science and human experience*. MIT Press, 1992.

22. Noë A. *Action in perception*. MIT Press, 2006.

23. Kant I. *Critique of the power of judgment*. Cambridge University Press, 2000.

24. Levinson J. Hume's standard of taste: the real problem. *J Aesthet Art Crit* 2002; 60(3): 227–238.

25. Frijda NH. The place of appraisal in emotion. *Cogn Emot* 1993; 7: 357–387.

26. Frijda NH. *The laws of emotion*. Lawrence Erlbaum Association Pub, 2007.

27. Scherer KR. The dynamic architecture of emotion: evidence for the component process model. *Cognit Emot* 2009; 23(7): 1307–1351.

28. Mueleman B and Scherer KR. Nonlinear appraisal modeling: an application of machine learning to the study of emotion production. *IEEE Trans Affect Comput* 2013; 4(4): 398–411.

29. Lazarus R. *Emotion and adaptation*. Oxford UP, 1991.

30. Lazarus R. *Stress and emotion: a new synthesis*. Springer Publishing Company, 2006.

31. Smith L. Piaget's developmental epistemology. In: Mueller U, et al. (eds) *The Cambridge companion to Piaget*. Cambridge Univ Press, 2009, pp. 64–93.

32. Meltzoff AN. Elements of a developmental theory of imitation. In: Meltzoff AN and Prinz W (eds) *The imitative mind: development, evolution, and brain bases*. Cambridge Univ Press, 2002, pp. 19–41.

33. Baron-Cohen S. *Mindblindness: an essay on autism and theory of mind*. MIT Press, 1997.

34. Robobuilder RQ-Titan. http://www.robobuilder.net/ (accessed 11 November 2015).

35. Ekman P and Friesen WV. Constants across cultures in the face and emotion. *J Pers Soc Psychol* 1971; 17(2): 124–129.

36. Ekman P, Freisen WV, and Ancoli S. Facial signs of emotional experience. *J Pers Soc Psychol* 1980; 39(6): 1125–1134.

37. Abrosoft FantaMorph. http://www.fantamorph.com (accessed 11 November 2015).

38. de Meijer M. The contribution of general features of body movement to the attribution of emotions. *J Nonverbal Behav* 1989; 13(4): 247–268.

39. Montepare JM, Goldstein SB, and Clausen A. The identification of emotions from gait information. *J Nonverbal Behav* 1987; 11(1): 33–42.

40. Bianchi-Berthouzeand N and Kleinsmith A. A categorical approach to affective gesture recognition. *Connect Sci* 2003; 15(4): 259–269.

41. Cortes C and Vapnik V. Support vector machine. *Mach Learn* 1995; 20(3): 273–297.

42. Mita T, Kaneko T and Hori O. Joint Haar-like features for face detection. In: *ICCV 2005, 10th IEEE International Conference on Computer Vision*, Beijing, China, Vol. 2, 15–21 October 2005. IEEE.

43. Fisher RA. On the interpretation of χ^2 from contingency tables, and the calculation of P. *J Royal Stat Soc* 1922; 85(1): 87–94.