OUKA
The University of Osaka
Institutional Knowledge Archive

Title	Designing Baby-Like Robots for Emotional Interaction with Older Adults
Author(s)	馮, 時
Citation	大阪大学, 2025, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/105267
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

Designing Baby-Like Robots for Emotional Interaction with Older Adults

SHI FENG

NOVEMBER 2025

Designing Baby-Like Robots for Emotional Interaction with Older Adults

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

BY

SHI FENG

NOVEMBER 2025

ABSTRACT

In the society where older adults take up more percentages in the populations, elderly care becomes more important. With the realization of the importance of elderly care and the investment of more infrastructures for older people such as caring facilities, and equipment improves for the better experience for older people, problems still exist. Though therapy, exercise and medication develop fast to make sure of the physical health of older people, the mental care of them is still not well addressed. Behavioral and Psychological Symptoms of Dementia (BPSD) and loneliness are the main issues harming the mental health and even the physical health of older adults in the caring facilities. Furthermore, the lack of care staff makes the situation worse with even less people accompanying older adults, and more burdens on the existing care staff. With more older adults admitted in the caring facility, and their length of life becoming longer, the mental care of them needs to be addressed more. Among the methods of mental care, the use of socially assistive robots is a method drawing more attention and possibilities. Using robots in the caring process can solve mental care and the lack of labor at the same time. Within social assistive robots, baby-like robots are becoming increasingly relevant in this industry. Studies have shown that the image of an infant has the best acceptance among older adults, and the caregiving behaviors older adults would perform on the baby-like robots can activate their neuron activities to prevent dementia as well as reduce their sense of loneliness. Furthermore, the emotional interaction conducted by baby-like robots makes them special among other robots. Until now, many baby-like

robots have been proven useful and effective in the mental care of older adults and have been used both in Japan and across the globe. However, there are still many about baby-like robots we do not understand and there can be more potential in baby-like robots to be discovered. Though baby-like robots showed usefulness, it is not understood what features in the design appealed to the older adults, and where in the older adult's psychology was triggered during the interaction with baby-like robots. In this thesis, we use Hiro-chan as the representative of baby-like robots. We focus on the three aspects of design in Hiro-chan, which are the morphological design, the functional design and the purpose design, referring to the designs of infant features, emotional functions, and the pursuit of the purpose of Hiro-chan towards the group interaction. We built a model for the multimodal infant features in baby-like robots in the morphological infant design of baby-like robots, discovered the effects of different emotions of Hiro-chan on the older adults and finally modified the purpose of Hiro-chan into a mediator to facilitate communications between older adults. By conducting this study on baby-like robots, we were able to find theoretical evidence for the effectiveness of baby-like robots and saw the potential of building a society that contains human-centered human-robot interaction by using baby-like robots as a media. As an extension to this thesis, a more accurate model of the relationships between multimodal infant features, baby schema and caregiving behavior can be formed using different infant modals, and finally have the model quantified with more detailed alterations in the infant models. Furthermore, this study has the potential to be applied not only to older adults for its effects in emotional interaction, but also to other people with difficulties in cognitive interactions, thus discovering a new path of socialization and building of relationships through emotional interaction to cognitive interaction.

TABLE OF CONTENTS

TABLE OF CONTENTS 7

1 INTRODUCTION **12**

1.1 Background 12

1.1.1 Aging society and caring facilities 12

1.1.2 BPSD and Loneliness of older adults 13

1.2 Related work 14

1.2.1 Socially assistive robots for elderly care 14

1.2.2 Baby-like robots 16

1.2.3 Minimal design of robots 20

1.3 Research Objectives 22

1.4 Contributions 24

1.5 Structure of Thesis 26

2 MORPHOLOGICAL DESIGN OF BABY-LIKE ROBOTS IN YOUNGER ADULTS **27**

2.1 Theoratical framework 27

2.1.1 Morphological design of Hiro-chan 27

2.1.2 Baby Schema 28

2.1.3 Caregiving Behavior 29

2.1.4 Design of Human Avatar 31

2.1.5 Current Study 31

2.2 Experiment 33

2.2.1 Participants 34

2.2.2 Experimental Design 36
2.2.3 Procedure 37
2.2.4 Data Processing 38
2.2.5 Results 40
2.2.6 Discussion 49

3 MORPHOLOGICAL DESIGN OF BABY-LIKE ROBOTS WITH OLDER ADULTS 60

3.1 Current study 60
3.2 Experiment 61
3.2.1 Participants 61
3.2.2 Experimental design 62
3.2.3 Hypothesis 62
3.2.4 Procedure 63
3.2.5 Data processing 64
3.2.6 Results 64
3.2.7 Discussion 73
3.2.8 Implications 73

4 FUNCTIONAL DESIGN OF HIRO-CHAN 78

4.1 Background 78
4.1.1 Effect infant emotions on adults 78
4.1.2 Emotional expressions in baby-like robots 79
4.1.3 Voice functions of baby-like robots 80
4.1.4 Passive social in socially assistive robots 81
4.2 Current study 82

4.2.1 Predictions 82
4.3 Experiment I 84
4.3.1 Participants 84
4.3.2 Experimental setup 85
4.3.3 Measurement 87
4.3.4 Procedure 88
4.3.5 Results 89
4.3.6 Discussion of experiment I 90
4.4 Experiment II 92
4.4.1 Participants 93
4.4.2 Experimental setup 93
4.4.3 Measurement 95
4.4.4 Procedure 95
4.4.5 Results 96
4.4.6 Comparison between experiment I & II 98
4.4.7 Discussion 99
4.5 General Discussion 101
4.5.1 Implications 102
4.5.2 Limitation 103
4.6 Conclusion 105

5 PURPOSE DESIGN OF HIRO-CHAN 106

5.1 Background 107
5.1.1 Interactive social 107
5.1.2 Balance theory 107

5.1.3 Social capabilities and social anxiety & social self-esteem . 108
5.2 Current Study . . . 109
5.2.1 Hypothesis . . . 109
5.2.2 Participants . . . 110
5.2.3 Apprautus . . . 111
5.2.4 Measurement . . . 111
5.2.5 Analysis . . . 114
5.3 Results . . . 115
5.4 Discussion . . . 118
5.4.1 Implications . . . 118
5.4.2 Differences in emotions of Hiro-chans . . . 119
5.4.3 Differences in gender . . . 120
5.4.4 Differences in social self-esteem level . . . 120
5.4.5 Limitations . . . 121
5.4.6 Conclusion . . . 122

6 GENERAL DISCUSSION AND CONCLUSIONS 124
6.1 Discussion . . . 124
6.1.1 Practical significance and design guidance . . . 124
6.1.2 Methodological and theoretical contributions . . . 124
6.1.3 Application and future work . . . 125
6.2 Conclusion . . . 128

ACKNOWLEDGMENT 130

BIBLIOGRAPHY 133

LIST OF PUBLICATIONS **142**

CHAPTER 1

INTRODUCTION

1.1 Background

1.1.1 Aging society and caring facilities

In recent years, due to the advance of modern medicine, the lifespan of people became longer. In the meantime, with the decreasing birth rate in many countries, the aging society is forming across the globe. Statistics show that people over 65 took up 27.7% in the year 2017. The number is anticipated to go up to 38.4% in 2065. With the more aged people and less younger ones caring for them, families are choosing to reside the older adults that need care in the caring facilities, which provides the health and mental care that family homes may not be able to provide. Older adults admitted to the caring facilities, as well as the number of caring facilities are increasing at a high rate each year in Japan. With the professional medical equipment in the caring facilities, the physical health of older adults is better treated. From this perspective, the caring facility for older adults is an improvement in the aging society and also an increasing need. Meanwhile, with more people admitted, problems emerge in the current situation. First of all, the lack of labor force in the aging society also impacts on the caring facilities. In 2023, 70.3% of the caring facilities were reported to have short of staff. Not only does this staff shortage affect the caring quality of the residents, but it could also make the burdens on the current staff on the facilities heavier. At the same time, while physical health is treated well in those facilities, the mental health of the older adults is less addressed, but still im-

portant. Such mental issues include Behavioral and Psychological Symptoms of Dementia (BPSD) and loneliness.

1.1.2 BPSD and Loneliness of older adults

Among the mental health issues faced by older adults in care facilities, Behavioral and Psychological Symptoms of Dementia (BPSD) and loneliness are two of the most common and challenging problems. BPSD may cause a wide range of symptoms associated with dementia, such as agitation, aggression, hallucinations, delusions, depression, anxiety, sleep disturbances, and even lead to violent behaviors. These symptoms not only affect the well-being and dignity of the older adults, but also significantly increase the burden on the care staff (Cerejeira et al., 2012). Meanwhile, loneliness is another deeply concerning problem among the older adults in the care facilities, especially in the current situation where there are not enough staff to accompany the older people. Loneliness could lead to depression, less control on their emotion, trouble sleeping or even lead to physical conditions (Hawkley & Cacioppo, 2010). In the caring facilities where people are more isolated and reluctant to interact with each other, the loneliness problem is more severe. With this growing problem of BPSD and loneliness of older adults admitted in caring facilities, and the shortage of staff, effective and innovative methods are required to solve the problem.

1.2 Related work

1.2.1 Socially assistive robots for elderly care

As an automatic method helping older adults in facilities deal with mental issues, socially assistive robots merged in recent years. Unlike physically assistive robots, this kind of robot focuses on verbal or emotional interactions with the users, providing a sense of company and social link as a therapy. We researched on the papers on social assistive robots for elderly care on Web of Science, Scopus and PubMed databases, using the keyword ("robot*" AND ("care*" OR "nurs*") AND ("elderly" OR "older"), including only the type of Article. The result showed an increasing trend of related research since 2008, and a sharper increase in 2018. This revealed that research and needed for socially assistive

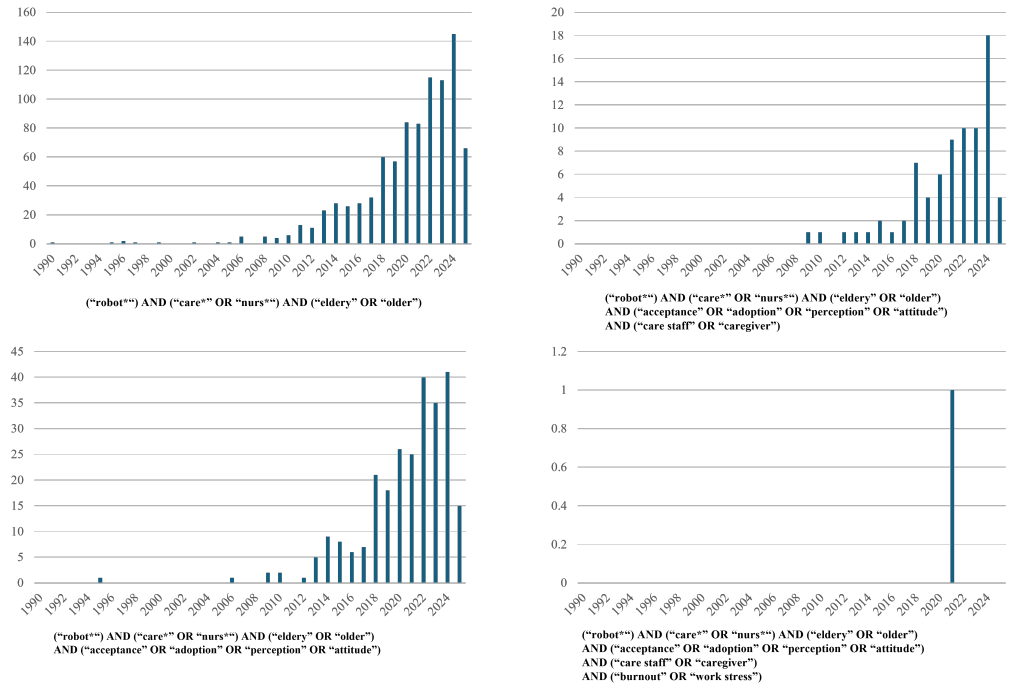


Figure 1.1: Search result for research papers in keywords concerning socially assistive robots and nursing homes

robots on elderly care have significantly increased over the past years. As an example, PARO, a robotic baby seal, is widely used in institutional care and has shown benefits in reducing stress and promoting social interaction among elderly residents (Wada et al., 2005). Other types of robots including Pepper[1], Babyloid[2], DarumaTO[3], Telenoid[4], Hiro-chan[5], have all been used for this purpose with different approaching methods, but all achieving good effects on the BPSD and loneliness with older adults. Within the socially assistive robots that exist nowadays, we categorized their design into two dimensions. The first is about the morphological design. We divided them on the basis of whether they are animal-like or human-like. The other about the interaction method. We divided them on the basis of cognitive interaction, which means the robot uses logical words to interact with the user, and emotional interaction, which stands for the robot communicating with the user by appealing to their emotions or building an emotional bond with the user. We believe that animal-like robots with cognitive interactions can be difficult for users to accept. While there are many robots with human-like designed and cognitive interactions as well as robots with animal-like designs and emotional interactions, robots with human-like design and emotional interactions can hardly be found. Human-like robots are believed to be able to make the user relate better, where emotional interactions require less cognitive ability in the user while consuming less cognitive burden. These traits make the combination of human-like design with emotional interactions good for elderly care. With very few robots having such combination, baby-like robots can fill this gap.

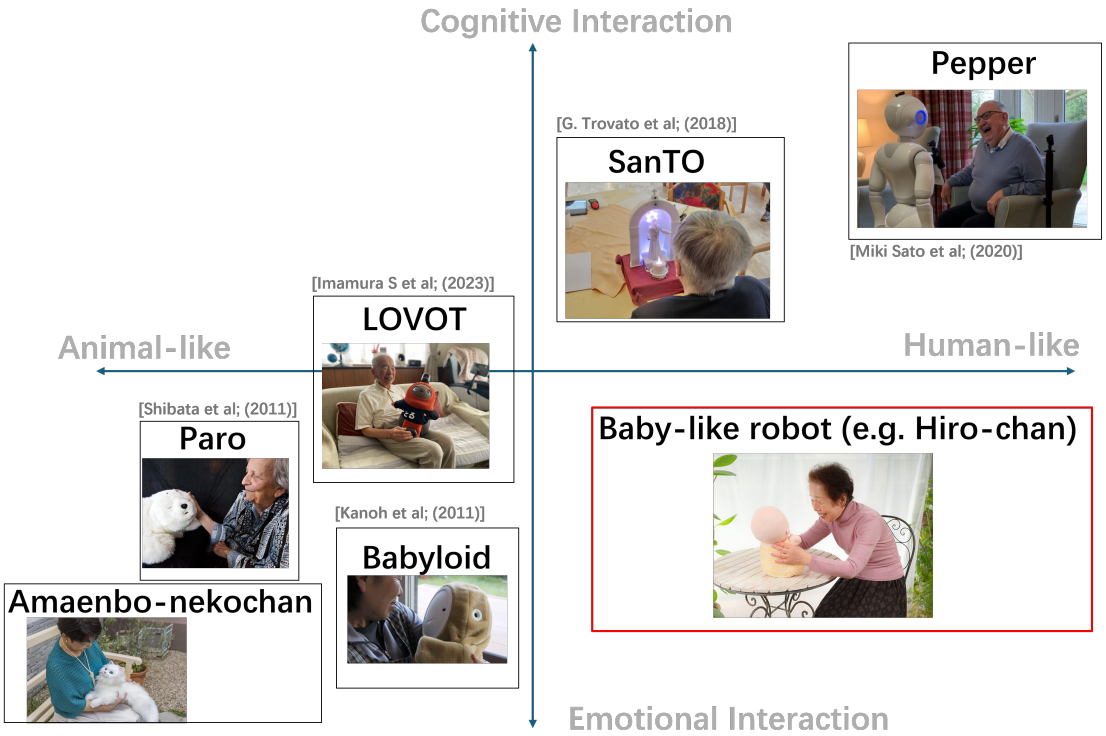


Figure 1.2: Socially assistive robots divided by morphological and interactive design

1.2.2 Baby-like robots

In recent years, baby-like robots have emerged as a novel category of socially assistive robots in the context of elderly care. Unlike service robots that aim to assist through physical manipulation, or other kinds of social robots that actively interact with the user, baby-like robots are primarily designed to elicit emotional engagement by evoking caregiving instincts in older adults. Their design draws on the principles of therapeutic presence, leveraging psychological mechanisms such as the baby schema effect, in which infantile features (e.g., large eyes, round face, soft body) instinctively trigger nurturing behavior in humans. The use of emotionally evocative designs in robotic therapy has gained substantial attention. Among these, PARO, as mentioned above, is

already widely used in caring facilities. While PARO mimics animal-like behavior, baby-like robots differentiate themselves by resembling human infants, targeting a deeper level of psychological association related to caregiving and parental bonding. Cohen-Mansfield et al. (2010) compared 25 kinds of stimuli to find out which are more positively engaged, and which are less positively engaged by older adults with dementia. The results show that real babies, real dogs and real humans are among the top-rated stimuli[6]. Among them, the human-human stimuli are the best while the baby is the second. This suggested that simulated social stimuli, such as dolls or robotic infants, are particularly effective among people with dementia. Though it is still not as good as interactions with other humans, it still surpasses other forms of non-pharmacological interventions in eliciting engagement. Other than this, in an experiment using a cat robot and baby-like robots, we found out that older adults perform more interactive actions for the purpose of caregiving on baby-like robots than the cat-like robot, while they interacted with cat-robot for the better feeling of themselves and the cat-robot. The interaction between baby-like robots and elderly users often resembles that of doll therapy, a practice where caregivers use infant dolls to reduce agitation and promote social behavior. Studies show that baby-like robots can elicit behaviors such as holding, talking, rocking, and nurturing — behaviors associated with enhanced emotional well-being and reduced caregiver burden (Mitchell et al., 2016). The underlying mechanism is thought to involve the activation of caregiving schemas in the brain, which fosters emotional bonding and a sense of responsibility, even in individuals with moderate cognitive impairment. Moreover, some baby robots, like Telenoid (Sumioka et al., 2014), use minimal information to encourage user imagination and projective empathy, where users attribute emotion and agency to the robot despite its lack

of expressiveness. One of the earliest prototypes in this field is Babyloid (Kanoh et al., 2011), a beluga-shaped robot capable of producing basic emotional expressions (e.g., joy, sadness) and responding to physical touch. Its design combines soft materials, infant-like audio cues, and simple actuations such as head movement and eye expressions. Babyloid has been used to reduce BPSD symptoms, showing temporary improvements in mood and increased verbal output from users. Another notable system is Hiro-chan, which represents a minimalist approach to baby-like robot design. Hiro-chan lacks facial features, using only infant-like recorded sounds to express different emotional states. It responds to tactile stimuli and internal state transitions but does not rely on visually rich expressions. In previous studies, Hiro-chan showed the same effect on older adults with or without a face [5]. This validated the idea of minimal design on Hiro-chan, proving the facial expression on Hiro-chan can be cut off. Another field research applied Hiro-chan to a caring facility for one month. While using the Hiro-chan with all functions including crying emotion for two months, negative impressions from both older adults and care staff were reported. After removing the crying from Hiro-chan and conducting the field research for another one month, most reviews were positive and participants remained the same positive opinions after one month of use. The results showed that impressions of older adults on Hiro-chan that do not cry stayed positive with no degression. Meanwhile the staff also claimed that residents of the caring facility became more emotionally stabilized and had less caring burden. This has proved Hiro-chan is capable of helping with mental care at a middle-ranged term and has the potential to be used stably in the long term.

While the Hiro-chan using minimal design is a promising method to relieve the situation for the older adults, there are still limitations to the current design.



Figure 1.3: Socially assistive robot Babyloid, adapted from Furuta, Y, et al, (2012)

Firstly, despite the seemingly obvious design of infants, the current morphosis design still lacks empirical backup, especially with the need for a minimal design. We understand based on the experimental result of Hiro-chan, that face is not necessary in the minimal design. However, it is still unclear to where the line of the effective minimal design draws. Baby-like robots lack guidelines for their morphological design. Secondly, the function of Hiro-chan was not thoroughly designed. Currently, the functions of Hiro-chan include recorded audios of real infants. However, from the experiment using Hiro-chan in the caring facility, users reported a 70% drop on the positive impressions when Hiro-chan includes crying function [7]. This suggests the effects of emotional expressions of Hiro-chan can largely affect the willingness to use the robot, user experience, effects after using, as well as the timespan of using the robot. We believe that controlling the audio of Hiro-chan can serve the users better. Moreover, since singular audio can make a difference on users, the combination of same or dif-



Figure 1.4: Socially assitive robot PARO, adapted from Kaoru Inoue et al, (2021)

ference audios could have more potential on influencing multiple aspects of the interaction. And yet, these functional designs have not yet been explored. Last but not least, we believe the target interaction method using Hiro-chan can be changed for better effects. According to Cohen-Mansfield, though baby faces are on the top list of stimuli, interactions with human are still the better stimuli [8]. Besides this, we believe provoking a human-human relationship can help with the boredom with robots, extending the timespan of usage for Hiro-chan in the long-term. With those points improved, Hiro-chan can serve as a better intermediate for mental care for older adults in the caring facilities.

1.2.3 Minimal design of robots

In recent years, an increasing number of studies have focused on minimal design in the field of human–robot interaction (HRI), which advocates simplifying

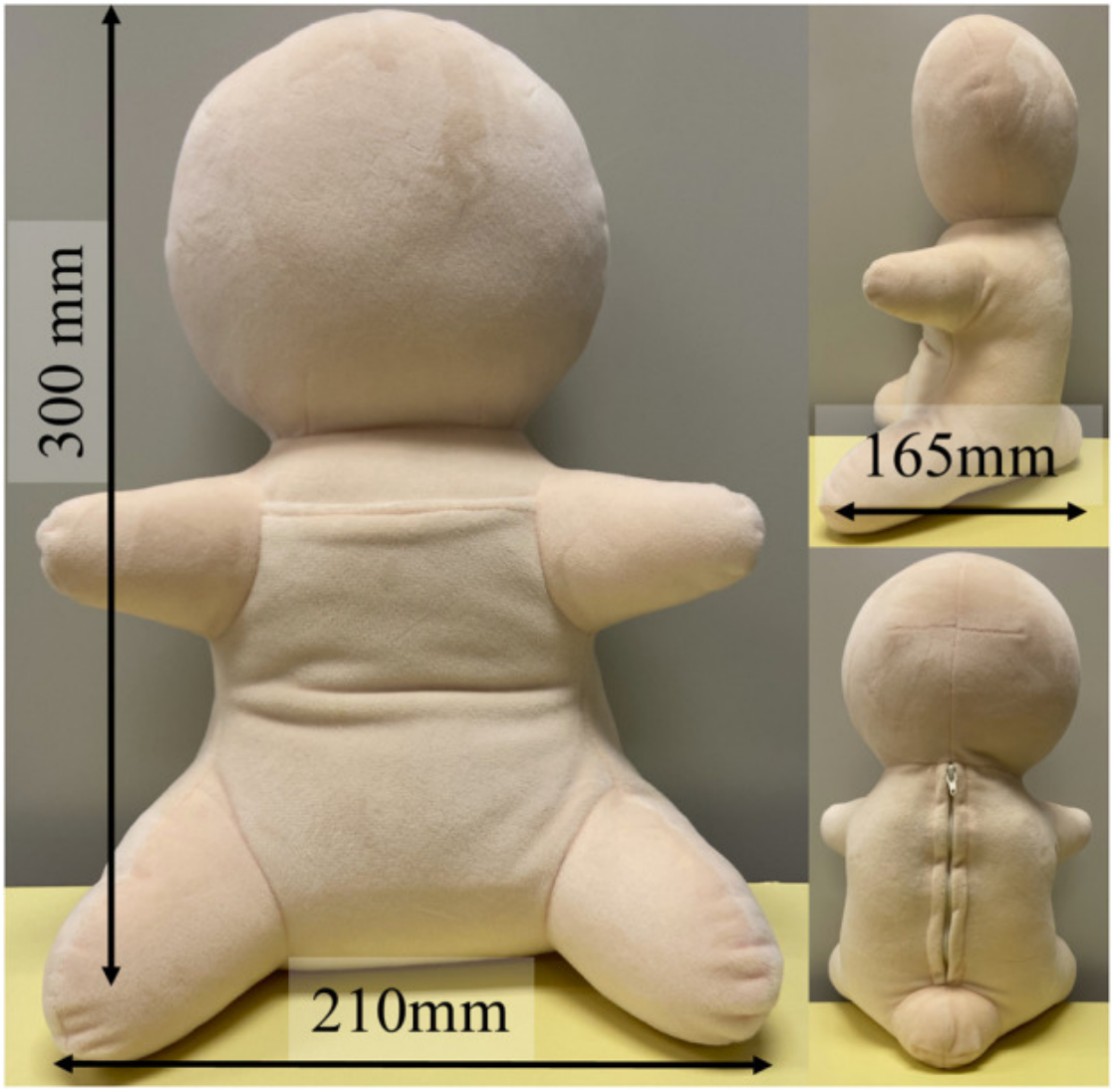


Figure 1.5: Socially assitive robot Hiro-chan, adapted from Hidenobu Sumioka et al, (2021)

robot appearance and expressiveness to its essential elements. Rather than enhancing realism through detailed features, minimal design aims to stimulate user imagination by removing unnecessary sensory and emotional cues. This aligns with the broader principle of subtraction design, where design choices are made not by adding complexity, but by intentionally eliminating non-essential components in order to clarify communication interfaces. This design philosophy has gained attention for its potential to reduce cognitive burden, partic-

ularly among users with sensory or cognitive limitations. On this aspect, the minimal design is useful in the interaction with older adults, especially those with dementia—often experience difficulty integrating information across multiple modalities. Therefore, simplifying robot outputs can enhance clarity and accessibility in interaction. Research has suggested that minimal cues, such as subtle body movements or vocal tone alone, may be sufficient to elicit empathy and engagement from users, especially when these cues are delivered in a controlled and consistent manner [9]. Minimal or abstract robot appearances also offer practical advantages. They reduce manufacturing costs, improve mechanical durability, and mitigate user discomfort related to the uncanny valley phenomenon. Moreover, users may project more personalized or context-specific interpretations onto robots with simplified forms, which can lead to deeper emotional engagement and longer-term acceptance. Such design has already seen effects in many robots such as Telenoid and Hiro-chan. Taken together, these studies suggest that minimal and subtraction-based design approaches offer not only aesthetic and practical benefits but also psychological and communicative advantages. These findings support the development of robots that are not necessarily human-like in form, but are nevertheless effective in fostering social connection, especially in contexts where emotional resonance is prioritized over functional complexity.

1.3 Research Objectives

In this research, we aim to understand the reasons behind the successful application of Hiro-chan in caring facilities and improve its functions so that it can deal with the loneliness and dementia of older adults better. To achieve this

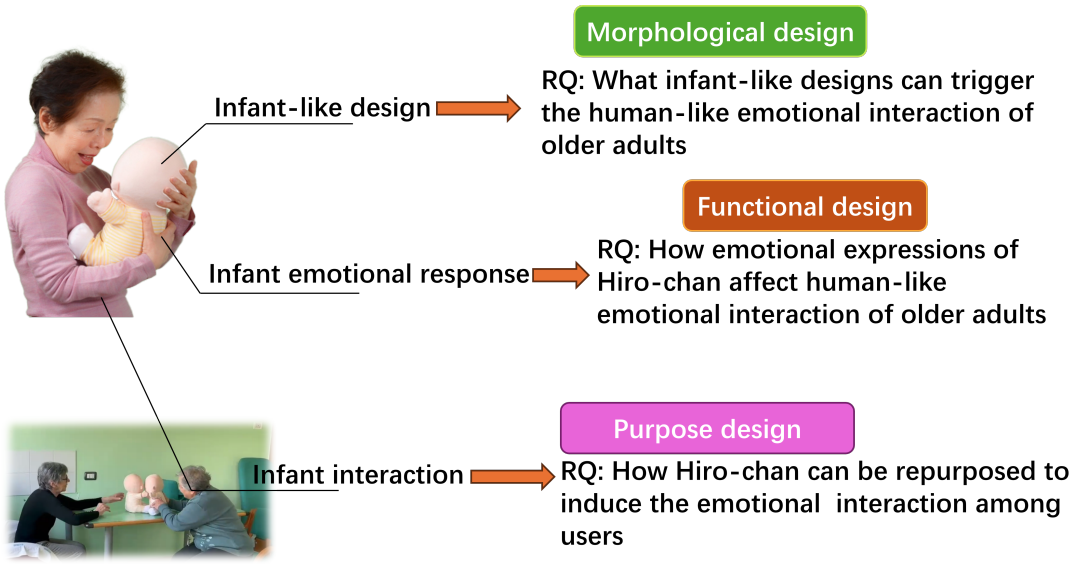


Figure 1.6: Designs on Hiro-chan of this thesis

objective, we divided the design of Hiro-chan into three steps to understand and improve them step by step. The first design is the morphological design. It refers to the infant features of Hiro-chan. Though morphology usually refers to the study of applying structure and form of animal or plants on products, in this thesis, we adopt a wider definition of morphology, which is the study and mimicking of characteristics of human infants, including all aspects of infants in the design, which includes form, voice, skin, temperature, smell, etc. We believe that baby-likeness is a main reason for the design of Hiro-chan and other baby-like robots. In the meantime, using too many infant features would cause downsides such as uncanny valley effect, increase in cost and more burden on the cognitive abilities of the users. That is why in this step, we aim to find the effective morphological design of Hiro-chan. The second step is the functional design. The current function of Hiro-chan only consists of the connection between accelerometer and the voices. In the experiment using Hiro-chan in the caring facility, we found a decrease in the impression and the use rate of Hiro-chan in Hiro-chans with crying function while no such decrease was spotted

in Hiro-chans that do not. Reports also showed cases of older adults rejecting care of crying Hiro-chans and the irritation of older adults dealing with crying Hiro-chans [7]. This showed us the current functional design of Hiro-chan can be flawed, and that we do not understand which emotional function contributes to the effective interaction with older adults and which one is harming the interaction. This is why we conducted the second step of functional design to understand what each emotional expression of the Hiro-chan voice function contributes to in the interaction. The third step is the interactive design. Since we have done the morphological design and the functional design of Hiro-chan, we have understood the reasons behind the successful implementation of Hiro-chan in caring facilities. The next step would be to utilize the findings for the improvement of interaction with older adults. In the previous experiment in the caring facility, communications among older adults and communications between older adults and care staff increased with the intervention of Hiro-chan. This is why we tested the ability of Hiro-chan to bring close the relationships of older adults, and its ability to increase willingness to communicate using the knowledge we learned in morphological design and functional design.

1.4 Contributions

We believe that with the understanding of psychology of older adults and the features in baby-like robots that appealed to these psychologies, we can provide guidelines for the effective design of baby-like robots. Furthermore, with the result of this study, we will be able to improve the current design of Hiro-chan to achieve better interaction or develop other methods of interactions with older adults that would ease their solitude and dementia. The first contribution of

this thesis is the design guidelines for baby-like robots. With understanding of the designs and functions in baby-like robots that can successfully build emotional bonds with older adults and trigger desired behaviors from them, and the unsuccessful designs that either triggered undesired behaviors or negative emotions in older adults, we earned a general guideline for baby-like robots in both morphological design and functional design.

The second contribution is the understanding psychology of older adults that baby-like robots appealed to. In this thesis, we found out the psychology behind older adults that made them like and willing to actively interact with baby-like robots. This understanding can further lead to multivariate elderly care by appealing to the same psychology.

The third contribution is that this thesis extended the interaction methods of socially assistive robots by facilitating the communication between older adults using Hiro-chan. This thesis pushed forward the study on group interaction among older adults using baby-like robots as a media. In the thesis, the emotional functions that facilitate communications, the target audience that are easier influenced by the existence of baby-like robots were discussed.

Finally, all the contributions above are to build a better environment in the aging society. This thesis showed how to improve the elderly mental care in the way of resolving loneliness as well as BPSD in elderly groups. The use and improvement of baby-like robots in this thesis can also lead to the relief of lack of human resources in caring industry, as well as the relief of heavy burden on the care staff in facilities.

1.5 Structure of Thesis

This thesis is structured as follows: Chapter 2 presents the study of morphological design of baby-like robots with younger adults. Chapter 3 shows the morphological design of baby-like robots with older adults. Chapter 4 shows the functional design of baby-like robot Hiro-chan which focus on the emotional expressions performing in different ways for older adults. In Chapter 5, we utilize the understandings from Chapter 2, 3 and 4 to further explore the purpose design of Hiro-chan to facilitate group interactions between older adults using Hiro-chan as a mediator and discusses the differences in the relationship and communications between the people involved caused by different emotions in Hiro-chan and the psychological difference in the participants. Chapter 6 includes the general discussion and findings of this thesis. The final chapter includes the general discussion, limitations and the future work of this thesis.

CHAPTER 2

MORPHOLOGICAL DESIGN OF BABY-LIKE ROBOTS IN YOUNGER ADULTS

This chapter discusses the morphological design of baby-like robots. It aims to understand how multimodal infant features in baby-like robots influence the psychology in older adults and relief their loneliness.

2.1 Theoratical framework

2.1.1 Morphological design of Hiro-chan

As mentioned in Chapter 1, the minimal design has already been applied and successfully utilized on baby-like robots for elderly care. However, there are still aspects of the design that have not been discussed for baby-like robots. First of all, though in general, baby-like robots' designs are based on the morphosis of infants, the design concepts of each baby-like robot are different. For instance, Babyloid made effort on the facial expression of the robot that resembles an infant with the shape like an infant in a cradle, whereas Hiro-chan aborted the facial expressions, but designing the robot to appear as an infant with torsos. There are also other baby-like robots such as Hinata-chan, which looks like a white baby doll with lined facial appearances. Though all the baby-like robots are designed to "look like a baby", there are no empirical evidence on where the design can be based on. There has not been any research on how the morphological design of baby-like robots should be performed for better effects, or how different morphological designs could affect the impression and usage of

baby-like robots. Secondly, though Hiro-chan robot successfully adopted minimal design, proving its efficiency and effectiveness on older adults, we still do not understand if the current design has any better adoptions of the minimal design. Which infant features are important to the design is still unclear. As mentioned above, guidelines towards the better morphological design of baby-like robots using minimal design are needed. In this chapter, we would like to discuss how to fill this gap and how to evaluate the effect of the design. In this chapter, we pursue a better morphological design on baby-like robots using minimal design and validate the design using empirical methods. We experimented with baby-like robots with different degrees of morphological design. We conducted the experiment with both younger adults and older adults.

2.1.2 Baby Schema

The concept of baby schema was first introduced by Konrad Lorenz (1943). Baby schema effect is described as a set of infant-like features that naturally trigger affection and caregiving behaviors in adults. These features usually include a large forehead, big eyes, and a small mouth, and are commonly seen in both human and animal infants (Kawaguchi et al., 2023). Research has shown that baby schema affects emotional, cognitive, and behavioral responses. People tend to prefer looking at infant faces rather than adult ones (Parsons et al., 2011), and this preference is found in both children and adults (Borgi et al., 2014; Van Duuren et al., 2003). Viewing baby-like images can also increase attention levels (Nittono et al., 2012) and cause people to smile more often (Hildebrandt & Fitzgerald, 1978). Recent studies in neuroscience have explored how baby schema influences brain and hormonal activity. Neuroimaging research shows

that baby schema can affect how people perceive and pay attention to visual information early in the process (Kringelbach et al., 2016; Luo et al., 2011). For example, seeing infant faces activates the nucleus accumbens in women without children, a brain area involved in motivation and reward (Glocker et al., 2009). Similar effects have been found in studies of infant sounds, especially crying. Hearing a baby cry activates brain regions related to emotion and caregiving, such as the amygdala, prefrontal cortex, and nucleus accumbens (Rilling, 2013; Swain et al., 2014; Witteman et al., 2019). These areas help increase emotional reactions, reduce irritation, and support long-term caregiving motivation. Hormonal changes also occur: crying raises oxytocin, which strengthens adults' sensitivity to babies, and also raises cortisol, which helps alertness but may lead to stress if prolonged. While many studies have examined how baby schema works in the brain or how to trigger it through facial features, fewer studies have focused on how baby schema actually leads to caregiving behavior. Most of the existing research is based on visual cues and less is known about how other signals—such as sounds or movements—might also play a role. Understanding these multimodal effects is important for designing systems that aim to encourage caregiving responses in humans.

2.1.3 Caregiving Behavior

Caregiving behavior usually refers to the actions that the parents take to protect and raise their children. The behaviors include tutoring for survival skills, playing with the children and feeding the children. At the same time, these behaviors have some notable traits that do not show in the behaviors towards adults. The traits include slower speed, exaggerated magnitude and higher pitch. This

kind of trait in the behavior is referred to as "infant-elicited social behaviors" by Stern (Stern, 1977). This infant-elicited social behavior is done by parents naturally to help with the cognitive learning process of infants and draw the attention of the infants. The "infant-elicited social behaviors" then can be broken down into two kinds, which are infant directed action and infant directed speech. infant directed action, or "motionese", present certain traits that appear naturally for the betterment of the infant. The traits include being more parsed, larger scaled and slower paced (Brand et al., 2002;Brand & Shallcross, 2008). The same goes for infant directed speech or "motherese", which is short, simple, and redundant (Durkin et al., 1982; Saint-Georges et al., 2013). It has been proved that motionese and motherese help with the growth of infants. A study by Fernald showed that four-months-old infants prefer motherese to normal speech towards adults (Fernald, 1985). Other studies also pointed out that motionese and motherese can be beneficial for object learning and language learning for infants (Koterba & Iverson, 2009; Nelson et al., 1989). Caregiving behaviors comes natural for humans, as the study of Rutherford & Przednowek suggested that both genders show similar adaptations to the motions towards infants (Rutherford & Przednowek, 2012). Not only is caregiving behavior beneficial for infants, it can also be soothing for adults. Both male and female experience change of level of hormones, which includes increased plasma oxytocin and reduced cortisol. Increased plasma oxytocin can lead to better social bonding and emotional attachment to the infant, where reduced cortisol level can reduce the stress. Lorenz believes that baby schema is one of the triggers for caregiving behavior in adults (Lorenz, 1943). That is why in this study, we utilize caregiving behaviors as an indicator of baby schema effect triggered by baby-like robots. On the other hand, though it is not an official categorization,

we use the term of non-caregiving behavior to refer to the behavior that does not treat the robots as an infant, which includes the carelessness towards the robot, and some violent behaviors towards the robot. We believe that baby schema is a critical effect that influences the impression on baby-like robots, and that the number of caregiving behaviors is both an important realistic reference to baby schema effect, and also a goal of a good morphological design.

2.1.4 Design of Human Avatar

Sumioka et al. studied about the development of human-like formed appearances in Japan history to give references for the design of human avatars. The research suggested that human-shaped bodies are more important than the facial expressions of human in the design of a human avatar for users to have a sensation of interacting with actual human. The research also pointed out that to achieve that, expressions of the head and the torsos need to be separated. Also, the expression of limbs is not important in the recognition of human-likeness in the avatar. The conclusions drawn in the research can provide some insight into the research for baby-like robots.

2.1.5 Current Study

In this study, we conducted experiments to empirically validate the better morphological design of baby-like robots. As the index of the effect of the morphological design, we adopt baby schema effect, which is shown by how many caregiving behaviors were conducted on the robots. Though we believe that

minimal design is applicable, that a certain degree of infant likeness would be enough for the robot to trigger baby schema, with the limited time and elements in the experimental environment, we think the cognitive load is not too big for the participants, and that with more infant features in the design, participants will see and treat the robots as a real infant, thus be more caring for the robot. That is why we made the following hypothesis H1 With more infant features in the morphological design, the more caregiving behavior the participants will perform on the baby-like robot. Likewise, we believe that the opposite is also true, where participants will not see the robots as infants when the infant features are not obvious, resulting in the performance of the more careless, even violent behaviors. This is why we hypothesize that H2 With less infant features in the morphological design, the more non-caregiving behavior the participants will perform on the baby-like robot. At the same time, we think it would be interesting to study the difference on performing caregiving behavior with 1) male and female participants; 2) participants with and without experience raising or playing with children. Though previous studies on real infants have concluded that mothers and fathers show similar adaptive behavior towards infants (Rutherford & Przednowek, 2012). In order to understand the above topics fully, applying baby-like robots and minimal design concept on a wider aspect, we decided to conduct the experiment on both younger crowds and older crowds. In this way, we would be able to understand not only the characteristics of baby-like robots, but also the design appealing to each age group. The experiment conducted on younger adults is experiment I, where the experiment on older adults is experiment II. We will cover the apparatus, participant group, process of the experiments, evaluation, the results and the discussion in the following parts of this chapter.

2.2 Experiment

Apparatus

In order to 1) find out how infant likeness affects the impressions on the users; 2) understand the borders of minimal design and 3) empirically validate the morphological design of baby-like robots, we prepared robots of different levels of morphosis design. Since we were to test the effect of minimal design, and it is difficult to analyze all the infant likeness elements, we focused on the elements that the current baby-like robots applied, which are the shape and the voice of infants. For this purpose, we prepared five robots with different morphosis shapes. The robots include baby-form, snow-form, cross-form, square-from and round-form. The baby-form is a form with all limbs of an infant and a clear illustration of a head. Cross-from is designed like a cross, with longer parts underneath than above. The purpose of the design is to make the robot have shape of the torso of an adult. Snow-form is designed to have a clear shape of the torso of an infant. While square-form and round-form are the ones without any torso, with two different shapes as comparison. The choice of robot shapes was based on the conclusion drawn from the Dogu research mentioned above. We categorized the shapes into forms with full infantile design, forms with full expression of an infant, the ones with torsos of an adult, the ones with torsos of an infant and the ones without any above, using two different shapes (round and square). Though it is not a fully accurate illustration of levels of morphological design of baby-like robots, we were hoping that this could provide us with some reference. For the same reason, we exclude the facial parts of infants such as eyes or mouth. It is proved in both the research on Dogu and minimal design, that facial expressions have little impact on the recognition of the figure,

especially for the older adults with dementia. Other than the shape design, we also tried to change the voice of the robot to another variable. Since the voice function of baby-like robots are using recordings of real infants, and it is difficult to morph the voice to make it sound like infants at different levels, we used two sets of varied voice functions. One is voice on, and the other is off. Other than the difference of shapes and voice, we managed to make the other variables of the robots as similar as possible. The designs we kept aligned include texture, function, weight and stuffing. All of the robots were made from plush fabric material with cotton stuffing inside. To make sure the stiffness felt the same, the cotton stuffing has similar amounts inside the robots. Within each of those robots, there is an ABS control module to control how the functions are triggered. The ABS module includes a 3-axis accelerometer for sensing the movement done on the robot, a speaker for making recorded infant voices, and a microprocessor using algorithms to trigger voice functions using accelerometer data. The algorithm is set to relate the acceleration of the sensors, which tells how active that the user is using the robot, to the voice functions. The more active the movements are, the happier the voice is for the robot. The emotions can be categorized into three levels, which are laughing, babbling and crying. With previous experience of older adults having bad impressions on crying baby-like robots, we decided to remove crying voice functions in this set of experiments, leaving only laughing and babbling voices.

2.2.1 Participants

For the experiment on younger adults, we recruited 45 subjects. 23 subjects are male and 22 are female. Their ages are between 20 and 50 (M =38, SD

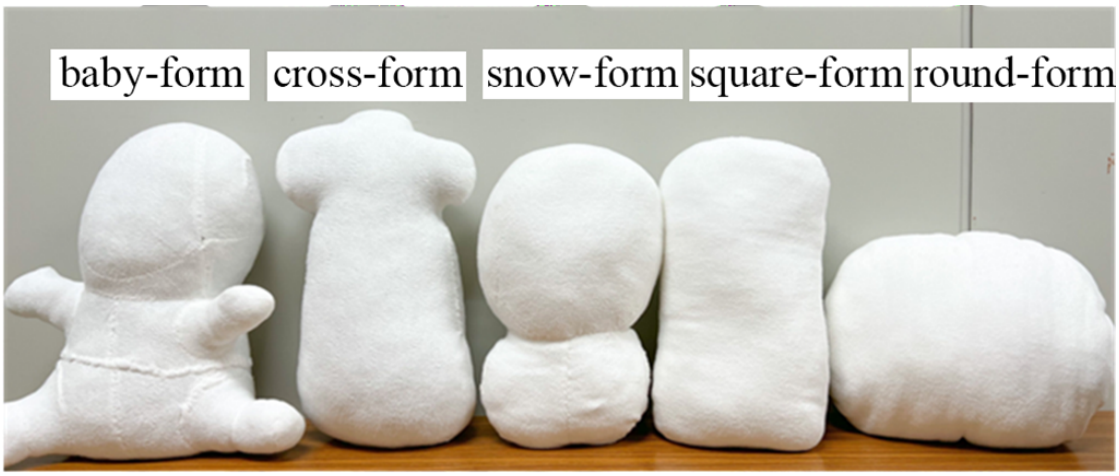


Figure 2.1: Five robots in different forms

=12). Since the experiment and the questionnaire, interview after the experiment was conducted in Japanese language, the recruits all have native Japanese level. Other than the age and the gender, there was no excluding process for the participants during recruiting. The sample size was decided using the reverse ANOVA analyzation to see how big a sample size could be effective in such experiments using MorePower 6.04 (Campbell & Thompson, 2012). Since in this experiment, we have three factors, which are the participant gender and robot voice as between-subject factor, robot form as within-subject factors, we assumed a three-factor ANOVA with a partial eta squared (η_p^2) of 0.14, significant level of 1% and 95% power. The equation for calculating Cohen's f as a first step is as below.

$$f = \sqrt{\frac{\eta^2}{1-\eta^2}}$$

In this case where η_p^2 =0.14, the calculation in this experiment is

$$f = \sqrt{\frac{\eta^2}{1-\eta^2}} = \sqrt{\frac{0.14}{1-0.14}} = \sqrt{\frac{0.14}{0.86}} \approx \sqrt{0.1628} \approx 0.403$$

After acquiring f, with the determination equation of non-centrality parameter by $\lambda = f^2 \cdot n$, where n is the sample size and λ is the non-centrality parameter.

With the calculation of non-central distribution in the MorePower software, we have a sample size of 40. Since the experience with children was a side research topic in this experiment, it was not set as a factor during the decision on the sample size.

2.2.2 Experimental Design

Since we are studying the different designs of robots, we decided to have all the participants experience all five forms of robots, with half of them experiencing robots with voice functions, the other half without. Thus, we divided the participants into two groups. One group with 25 participants, within in which 13 are males and 12 are females (M=39, SD=12), interacted with all five forms of robots with voice function on. This group is marked as Group V-on. The other group, which is Group V-off, consisting of 10 male participants and 10 female participants (M=36, SD=11), interacted with the five forms of robots with the voice function off. The experiment was designed to be conducted in a confined room, with the robot placed on the lap of the participant. We believe this could increase the intimacy with the robot so that we could possibly observe more actions done on the robots. To acquire the opinion on the robots from the participants, we let the participants interact with each robot form for about one minute. In order for the participants to interact with the robots freely without interference and get the natural data from the experiment, the participants were left alone during the interaction. Also, to eliminate the possibility that the order affecting the impression, during the experiment, the order of the five robots the participants interacted with were random. The data extracted from the experiment comprises two types. One is the questionnaire collected

from the participants. The other is the video recording of the experiment. The questionnaire consists of three ranking questions and one interview question.
1. Baby-likeness: Please rank the five robots in the order you found most baby-like.
2. Fun to play: Please rank the five robots in the order you found most fun to play with.
3. Easy to play: Please rank the five robots in the order found most easy to play with.
We asked the participants to give the above three rankings, and the reasons for the ranking. Note that we allowed the same ranking for different robot forms. Also, the reasons can be left blank if the participants failed to provide one. In the interview of the questionnaire, we asked the participants if they have played with children recently or if they have experience raising children. The rankings are for the statistical analysis in the experiment, while the reasons for the rankings and the interview were for reference. For the video recording of the experiment, we planned to observe the caregiving and non-caregiving behaviors from the participants. As stated before, we believe the two behaviors are the indicators for baby schema effect, thus be a hint for the effect of the morphological design of baby-like robots. With this experiment design, we hope to extract the information we need to validate our hypothesis and predictions.

2.2.3 Procedure

When the participant arrived, a consent was taken prior to the experiment. In the consent, we asked the participants for permission to use recordings of the experiment, with or without a censor to the participant's face. After the con-

sent, the participant was taken to the experiment room. The experimenter then briefed the participant about the experiment, that participant would be interacting with five different types of robots one by one for one minute each. And that the participant could interact with the robots freely. Then the experimenter asked the participant if he or she had any questions. After the explanation, the experimenter brings the first robot inside for the participant to interact with. Then the experimenter would leave the room for approximately 1 minute. After 1 minute, the experimenter would come in and collect the robot, bringing the next robot at the same time. The same process was repeated until the participant had interacted with all five robots. After the interaction, the experimenter reentered the room, bringing all five robots and placed them on the table. The order of the five robots were the same as the order they were brought in for interaction. The experimenter then asked the questions written on the questionnaire. After the questionnaire was finished, the experimenter saw the participant out and finished the experiment for this participant. All the other participants had the same procedure in the whole experiment.

2.2.4 Data Processing

In the experiment, we collected 1) rankings of robots on the three questions with reasons; 2) interview answers; 3) video recordings. Among those, 1) and 3) are for statistical analysis and 2) is for reference. In the following part, we will explain how each data collected is used and processed. Firstly, the rankings of robots are transferred into scores for a straightforward presentation of the results. If the robot form was placed in the first place, the score is 5. If it was placed in fifth place, the score then is 1. We collected and analyzed the scores

of all the robot forms. The reasons for the rankings were used to help us understand the results drawn. Interview results were used for analysis if experience with children influences the result. Finally, the recordings were used to analyze the caregiving and non-caregiving behaviors. After viewing the recordings of the experiment, we categorized the movements on the robots into more detailed ones for analysis. We broke down the caregiving behavior into "infant directed actions" and "infant directed speech". Whereas non-caregiving behavior was divided into "non-infant directed actions" and "non- infant directed speech". We believe this breakdown could help us understand reasons behind differences in behaviors, giving us deeper insight into the results. Unlike the categorizations for presentation of results, to be more specific when viewing the videos, the movements of the participants were categorized into different types. The types include exercise, contact, infant directed speech, non-contact, non-teasing, violent and non-infant directed speech. Within those, 1) exercise, which stands for the active but gentle interactions with the robots such as lifting the robots and swinging the robots in the arm; 2) contact, which stands for the soft touches and the pettings on the robot; 3) non-contact, which stands for actively responding to the robots' presence or actions, even without physically interacting with the robots, are categorized as the "infant directed actions". The "non-infant directed actions" include 4) non-teasing, which stands for the actions towards a machine rather than an infant such as checking the voice source, flipping the robots to verify the functions; 5) violent, which as the name suggests, are the more aggressive actions not meant for infants, such as throwing the robots in the air, harshly slaps the robots. The infant directed speech is as explained in the Caregiving behavior part, a slower paced, higher-toned speech. In this experiment, the speech that showed interest in the robots and including

contents meant for infants such as “have you eaten” or “what is your name” were also categorized as infant directed speech. The non-infant directed speech refers to the speech not meant for infants, which includes thinking out loud, speaking of the robots’ mechanisms and the intonation words like “uh-huh” or “ooh”. We performed four-way mixed ANOVA on the scores and the movements. The between-subject factors are participant gender, Groups V-on and V-off and experience with children. The within-subject factor is the robot forms. We conducted this ANOVA for 1) Baby-likeness score; 2) Fun to play score; 3) Easy to play score; 4) Caregiving behavior; 5) infant directed actions; 6) infant directed speech; 7) Non-caregiving behavior; 8) Non-infant directed actions; 9) Non-infant directed speech. To confirm an assumption of sphericity, we first applied Mauchly’s test for the within-subject factor of robot type. Degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity if the test failed. We conducted post-hoc comparisons with Bonferroni corrections to confirm the differences with each robot form in detail.

2.2.5 Results

Baby likeness

Fig. 2.3 shows the means and standard deviations of the baby-likeness scores for the robots with each form. Baby likeness had a main effect of the robot form ($F_{(4,148)} = 57.46$, $p < .001$, $\eta_p^2 = .19$). Post-hoc comparisons with Bonferroni corrections showed that the baby-form robot, which was rated as the most baby-like robot, had a significantly higher score ($M = 4.88$, $SD = .46$) than the robots with other forms. There were other significant differences between the cross

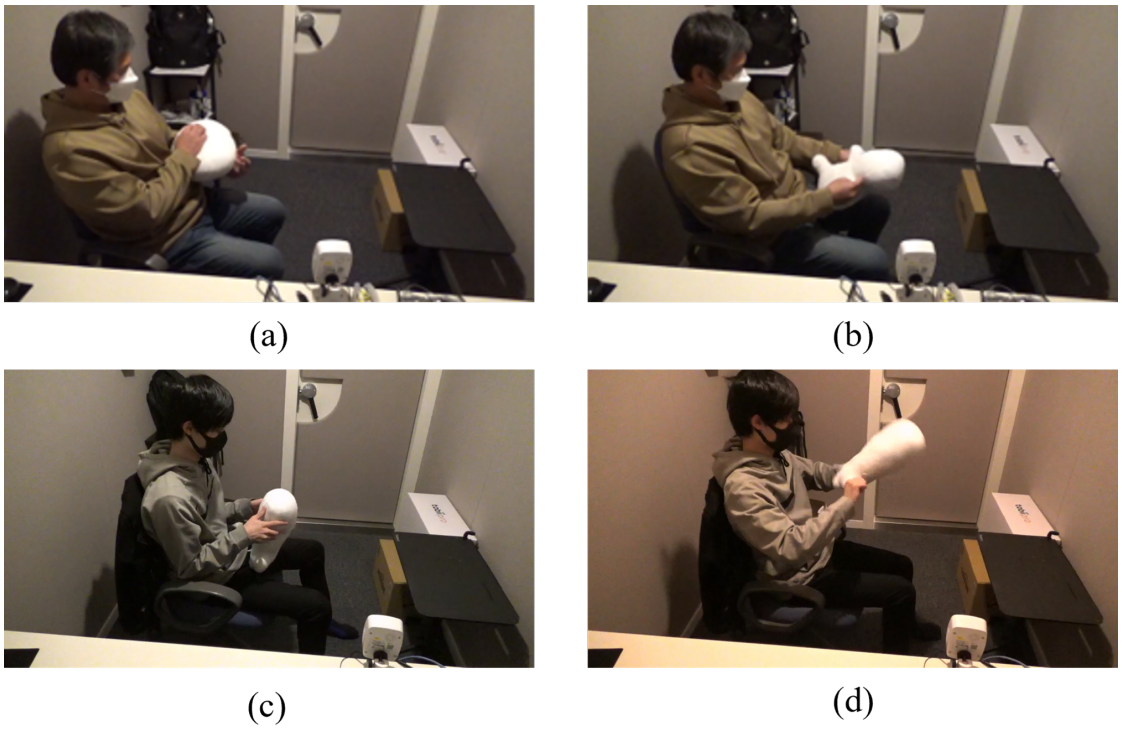


Figure 2.2: (a) Exercise action; (b) contact action; (c) non-teasing action; (d) violent action

($M = 2.71, SD = 1.21$) and square forms ($M = 1.85, SD = .91$) with $p = .029$, the snowman ($M = 3.51, SD = .87$) and square forms with $p < .001$, and the snowman and round forms ($M = 2.41,\ SD = .89$) with $p < .001$. There were also main effects of the voice group ($F_{(1,37)} = 7.114$, $p = .011$, $\eta_p^2 = .16$) and the gender group ($F_{(1,37)} = 7.114$, $p = .011$, $\eta_p^2 = .16$).

Fun to play with a robot

Fig. 2.4 shows the means and standard deviations of scores with each form concerning fun to play with a robot. We found a main effect of the robot form with ($F_{(4,148)} = 16.18$, $p < .001$, $\eta_p^2 = .30$). A post-hoc comparison with Bonferroni corrections for the robot form showed a similar tendency to baby likeness. The

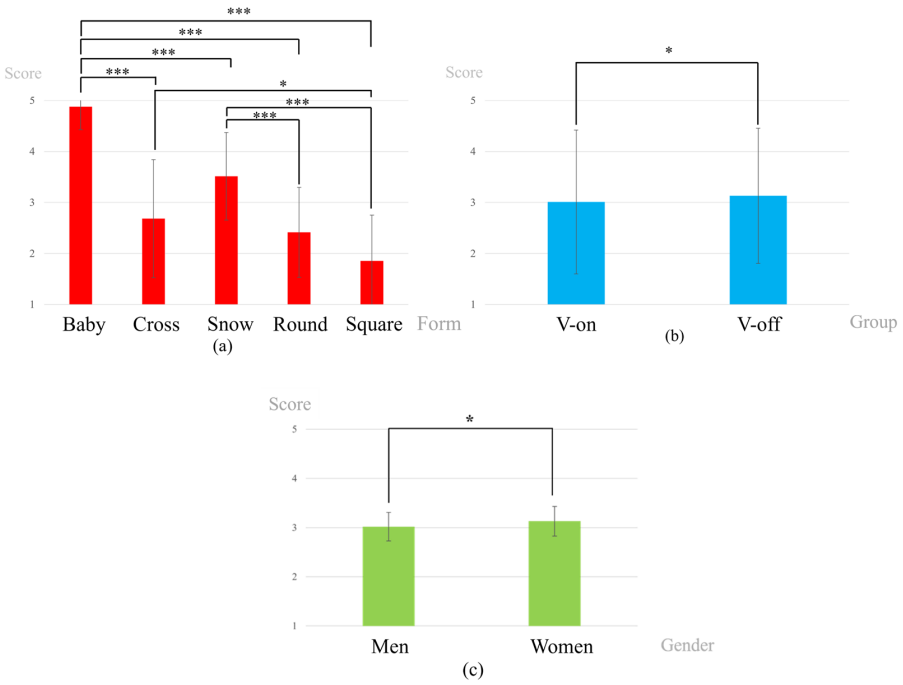


Figure 2.3: (a) Baby-likeness scores among all robot forms; (b) baby-likeness scores against robot voice; (c) baby-likeness scores against participant gender ($* : 0.01 < p < 0.05, ** : 0.001 < p < 0.01, * * * : p < 0.001$)

robot with the baby-form ($M = 4.39, SD = 1.05$) was significantly rated as the most fun to play with. There were also significant differences between the snow-man ($M = 3.24, SD = 1.18$) and square forms ($M = 1.20, SD = 1.23$) with $p < .001$. We also found a main effect of the robot voice ($F_(1, 37) = 11.76, p = .001, \eta_p^2 = .24$).

Easy to play with a robot

The scores in Fig. 2.5 show the ranking results for easy to play with, and a main effect was found only with robot form ($F_{(4,148)} = 57.46, p < .001, \eta_p^2 = .64$). A post-hoc comparison with Bonferroni corrections for the robot form showed the same tendency to fun to play: The robot with a baby-form ($M = 4.39, SD = 1.12$) was significantly rated as the most fun to play with. There was also a significant

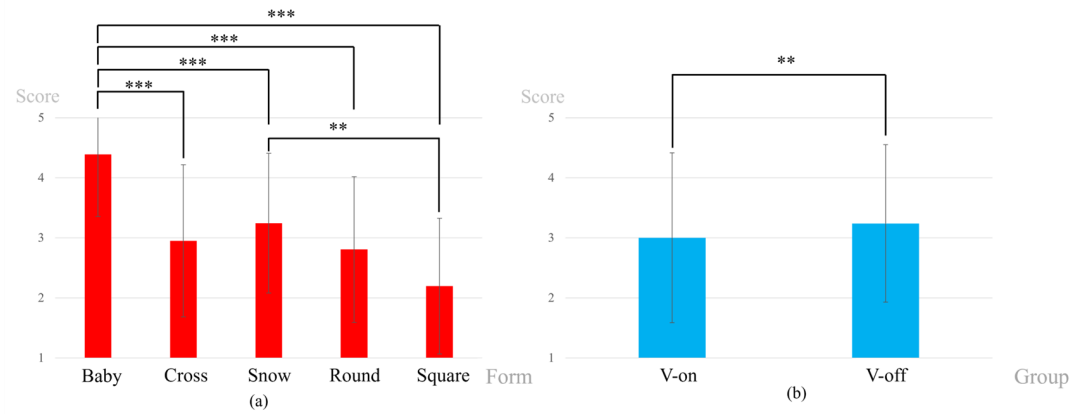


Figure 2.4: (a) Fun to play with scores among all robot forms and (b) fun to play with scores against robot voice
$(* : 0.01 < p < 0.05, ** : 0.001 < p < 0.01, *** : p < 0.001)$

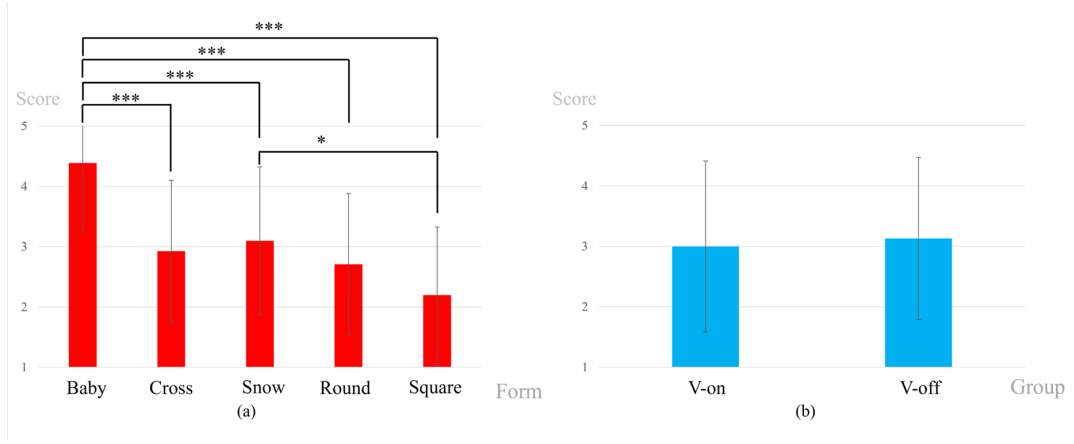


Figure 2.5: (a) Easy to play with scores among all robot forms and (b) easy to play with scores against robot voice
$(* : 0.01 < p < 0.05, ** : 0.001 < p < 0.01, *** : p < 0.001)$

difference between the snowman ($M = 3.10, SD = 1.23$) and square forms ($M = 2.20, SD = 1.13$) with $p = .032$.

Caregiving behavior

Total behaviors Fig. 2.6 shows the results of caregiving behaviors. We found a main effect behavior with the robot forms ($F_{(4,148)} = 9.19, p < .001, \eta_p^2 = .18$).

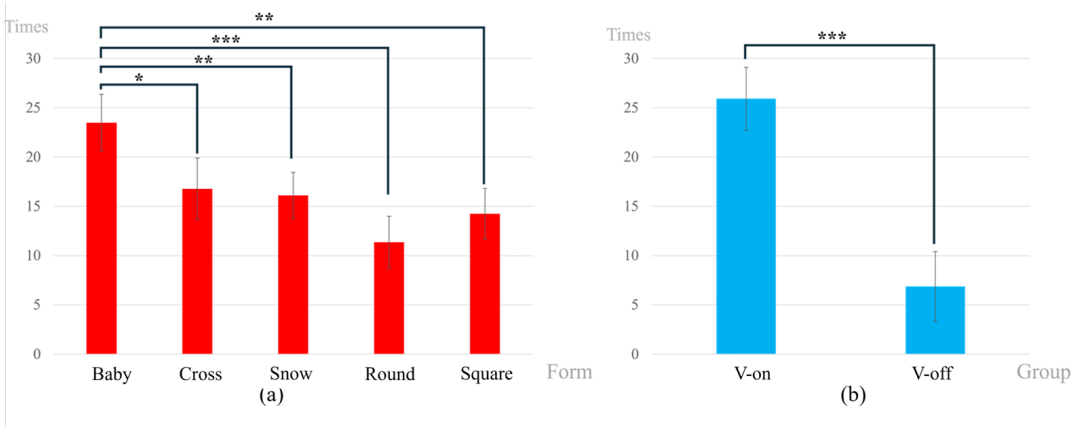


Figure 2.6: (a) Caregiving behavior among all robot forms and (b) caregiving behavior against robot voice
$(* : 0.01 < p < 0.05, ** : 0.001 < p < 0.01, *** : p < 0.001)$

A post-hoc comparison with Bonferroni corrections showed that caregiving behavior was observed most with the baby-form ($M = 23.49, SD = 2.87$), with a significant difference against all other forms. Participants performed more caregiving behaviors for the baby-form than the cross-form ($M = 16.78, SD = 3.73, p = .042$), the snowman-form ($M = 16.10, SD = 2.35, p = .004$), the round-form ($M = 11.35, SD = 2.64, p < .001$), and the square-form ($M = 14.24, SD = 2.58, p = .002$). There was also a main effect of the robot voice ($F_{(1,41)} = 15.99, p < .001, \eta_p^2 = .28$), which increased the caregiving behaviors.

Table 2.1: Means and standard deviations of Behaviors.

	Caregiving Behavior		Non-caregiving Behavior	
Mean (SD)	17.28 (21.04)		15.60 (12.95)	
	infant directed		Non-infant directed	
	actions	speech	actions	speech
Mean (SD)	13.31 (13.87)	3.96 (8.53)	13.96 (11.28)	1.64 (3.41)

Infant directed actions The main effect of the robot forms ($F_{(4,164)} = 11.83, p < .001, \eta_p^2 = .22$) on the infant directed actions was significant. A post-hoc comparison with Bonferroni corrections showed that infant directed actions were

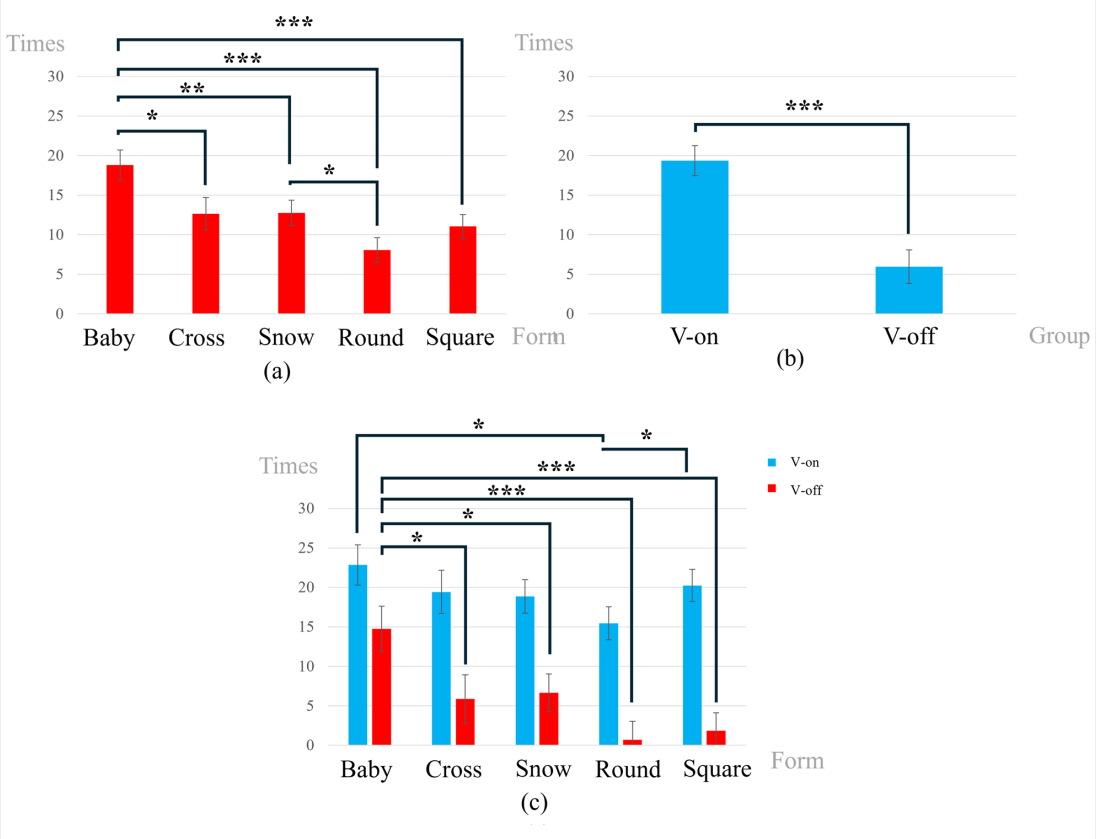


Figure 2.7: (a) infant directed actions among all robot forms; (b) infant directed actions against robot voice; (c) infant directed actions among robot form x robot voice
($* : 0.01 < p < 0.05, ** : 0.001 < p < 0.01, *** : p < 0.001$)

performed most on the baby-form ($M = 18.80, SD = 1.92$), with a significant difference against all other robot forms: cross-form ($M = 12.64, SD = 2.06, p = .012$), snow-form ($M = 12.75, SD = 1.60, p = .003$), round-form ($M = 8.08, SD = 1.56, p < .001$), and square-form ($M = 11.05, SD = 1.53, p < .001$) (Fig. 2.7(a)). We found no significant difference between other pairs of robot forms except for the snowman and round pair ($p = .016$). There was also a significant main effect for the robot voice on infant directed actions ($F_{(1,41)} = 22.32, p < .001$, $\eta_p^2 = .35$). Group V-on participants showed more infant directed actions (Fig. 2.7(b)). We found no significant effect on gender ($F_{(1,41)} = 2.24, p = .142$, $\eta_p^2 = .052$). There was a significant interaction between the robot form and the condi-

tions ($F_{(4,164)} = 2.58, p = .039, \eta_p^2 = .059$) (Fig. 2.7(c)). A post-hoc comparison with Bonferroni corrections showed that the robot's voice significantly influenced the number of infant directed actions induced by each form. Participants made significantly more infant directed actions to the baby-form than the other forms when interacting with the robots without a robot's voice (baby-cross: $p < .05$; baby-round: $p < .001$; baby-snowman: $p < .05$; baby-square: $p < .001$). They also performed significantly more infant directed actions to the snowman-form than the round-form (snow-round: $p < .05$). However, the influence of the robot's form weakened when the participants interacted with the robots that elicited an infant's voice. We found significant differences between the baby and round forms and the square and round forms, but no other combinations. We identified no other interaction effect among the robot forms, the voice conditions, and gender.

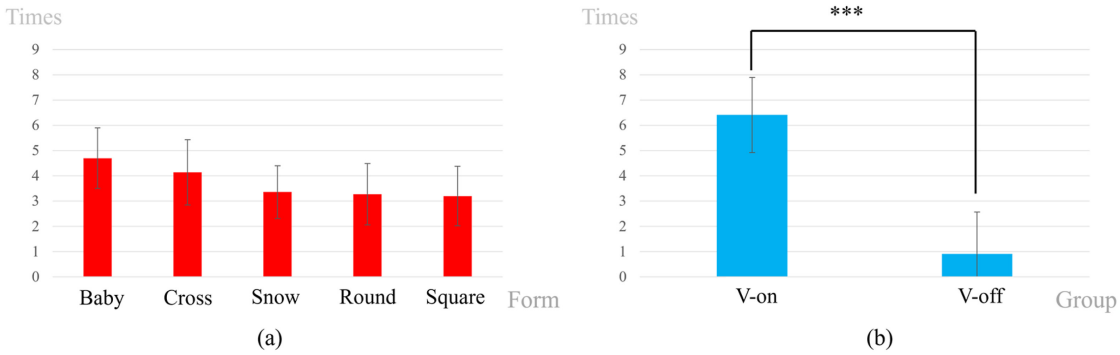


Figure 2.8: (a) infant directed speech among all robot forms; (b) infant directed speech against robot voice
$(* : 0.01 < p < 0.05, ** : 0.001 < p < 0.01, *** : p < 0.001)$

Infant directed speech The main effect that affects infant directed speech towards the baby-like robots was the robot's voice ($F_{(1,41)} = 11.25, p < .001, \eta_p^2 = .22$), which is shown on Fig. 2.8(b). There were no main effects in either the robot form or the gender of the participants (Fig. 2.8(a)).

Non-caregiving behavior

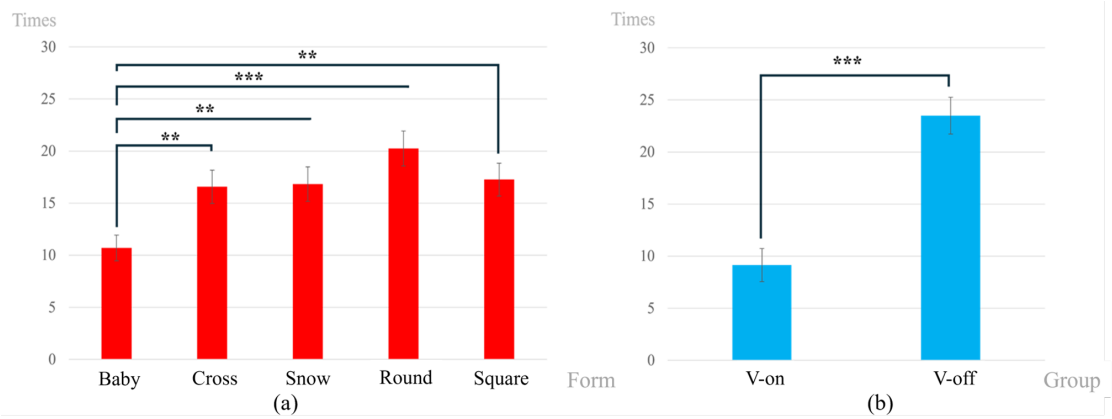


Figure 2.9: (a) Non-caregiving behavior among all robot forms and (b) non-caregiving behavior against robot voice
$(* : 0.01 < p < 0.05, ** : 0.001 < p < 0.01, *** : p < 0.001)$

Total behaviors Non-caregiving behaviors showed contrary results to the caregiving behaviors (Fig. 2.9). The main effects of the non-caregiving behaviors are the robot form ($F_{(4,164)} = 9.74, p < .001, \eta_p^2 = .19$). A post-hoc comparison with Bonferroni corrections showed that participants showed more non-caregiving behavior to the baby-form ($M = 10.70, SD = 1.22$) than all the other forms: cross-form ($M = 16.57, SD = 1.59, p = .003$), snow-form ($M = 16.82, SD = 1.65, p = .001$), round-form ($M = 20.24, SD = 1.67, p < .001$), and square-form ($M = 17.26, SD = 1.58, p = .003$). We found a main effect of the robot voice ($F_{(1,41)} = 36.26, p < .001, \eta_p^2 = .47$) and the interactions of the robot form, the robot voice, and gender ($F_{(4,164)} = 2.97, p = .021, \eta_p^2 = .68$). A post-hoc comparison with Bonferroni corrections showed that in Group V-on, there were significant differences in the male participants with the baby-form ($M = 5.85, SD = 2.26$) against the snow-form ($M = 13.39, SD = 3.06, p = .050$) and the round-form ($M = 17.92, SD = 3.09, p = .005$). There were no significant differences among the female participants. In Group V-off, there were significant differences in the

male participants between the baby-form ($M = 18.50, SD = 2.26$) and the snow-form ($M = 27.20, SD = 3.49, p = .046$). In the female participants, significant differences were observed with the baby-form ($M = 13.10, SD = 2.58$) against the snow-form ($M = 22.70, SD = 3.49, p = .019$, the round-form ($M = 30.20, SD = 3.52, p < .001$), and the square-form ($M = 27.20, SD = 3.32, p = .002$).

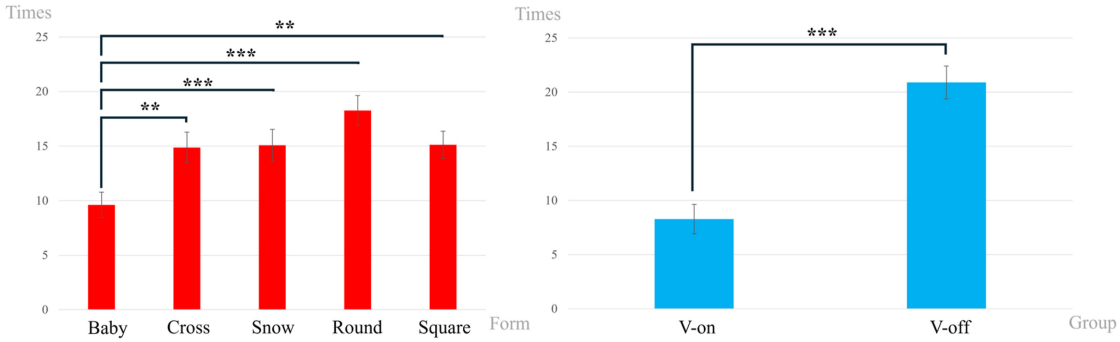


Figure 2.10: (a) Non-infant directed actions among all robot forms and (b) non-infant directed actions against robot's voice
($* : 0.01 < p < 0.05, ** : 0.001 < p < 0.01, *** : p < 0.001$)

Non-infant directed actions The main effects of the non-infant directed actions that have significant differences are the robot's form ($F_{(4,164)} = 10.22, p < .001, \eta_p^2 = .20$) and its voice ($F_{(1,41)} = 38.39, p < .001, \eta_p^2 = .48$). We found no interactions between either of these factors. Post-hoc comparison found that non-infant directed actions are done significantly less on baby-form ($M = 9.62, SD = 1.17$) than all other forms: cross-form ($M = 14.87, SD = 1.42, p = .004$), snow-form ($M = 15.08, SD = 1.47, p < .001$), round-form ($M = 18.27, SD = 1.38, p < .001$) and square-form ($M = 15.12, SD = 1.25, p = .003$) (Fig. 2.10).

Non-infant directed speech The main effect of the non-infant directed speech was found for the robot's voice ($F_{(1,41)} = 16.96, p < .001, \eta_p^2 = .29$), which is shown on Fig. 2.11(b). There were no main effects in either the robot form or the gender

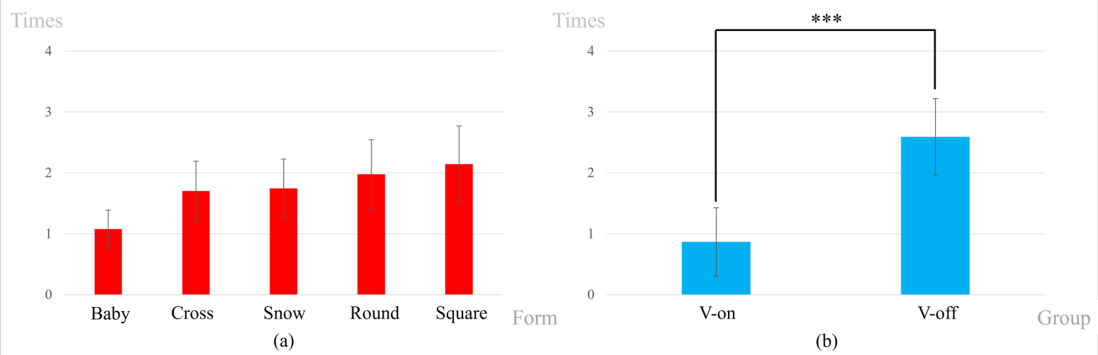


Figure 2.11: (a) Non-infant directed speech among all robot forms and (b) non-infant directed speech against robot's voice
($* : 0.01 < p < 0.05, ** : 0.001 < p < 0.01, * * * : p < 0.001$)

of the participants (Fig. 2.11).

2.2.6 Discussion

Implications

In this experiment, we used different degrees of infant likeness in the morphological design in baby-like robots to test its influence on the effect of the robots, which is expressed as caregiving behavior done on the robots. The questionnaire results showed the impression on the robots. It showed that participants think the baby-form being most baby-like. The second most baby-like form, which is snow-form, is also significantly more baby-like than the least baby-like robot form, the square form. The most and second most babylike forms also are significantly different. This enlightened us on the morphological design principles on the infant factors that would make the robot more baby-like. Fron the results, we concluded that the designs with a general shape of the torso of the infant is important in the illustration of baby-likeness. The snow-form with a torso

of the infant is significantly more baby-like than those that do not, including the round and square-forms. The cross-form, which has an illustration of an adult's torso, was only significantly more baby-like than square-form. The baby-form, which has the display of both torso and general shape, is the most baby-like of them all. Meanwhile, for robot forms that are easy to play with and fun to play with, baby-form is still significantly better than the rest of the forms, and snow-form is not better than round-form, unlike in baby-likeness. We found explanations for this in the post-experiment interview session, where the participants answered that the round-form resembled the shape of a ball, thus the higher scores easy and fun to play. We also found that for baby-likeness, main effect exists for gender. Since there is also main effect for voice on/off group, we further analyzed the situations with voice and gender and found that for female participants in the Voice-On group, there are multiple cases where many robot forms were marked with the same ranking. This may have caused the baby-likeness to have inflated scores and increased the scores for baby-likeness overall. Also, this suggested that the morphological design of infant features may appear similar for female when the voice is on. While it seems obvious that the more infant features (torso, general shape) the robots have, the more baby-like it becomes, from the results of behavioral analysis, we discovered that the caregiving behavior is not simply related to the baby-likeness. For caregiving behavior, both the baby-like form and voice had significant influence on the number of caregiving behaviors. Caregiving behavior is done on baby-form, which has both the general shape of an infant and the illustration of limbs, significantly more than all other forms. In the meantime, we found no significant difference between the robot forms that have an infant's torso (snow-form), an adult's torso (cross-form) and the ones that do not illustrate

any torso (round-form, square-form), which is different from the questionnaire results of baby-likeness. This result shows that the baby-like general shape, as well as an infant torso, though abstract, could be important in triggering caregiving behavior. We found significant differences in groups V-on and V-off, but no main effect in the interaction between robot form and robot voice, which means that having an infant voice can increase caregiving behavior regardless of how the robot is shaped. The results above support hypothesis H1. The results for non-caregiving behavior are the same ass caregiving behavior, where there were significant main effects for robot form and robot voice. The further analysis of post-hoc gave the same results as well. Non-caregiving behavior was significantly done less on robot-form than all other forms. Group Voice-on had significantly less non-caregiving behaviors than Voice-off, while main effect was not found in interaction between form and voice. This indicates that having an infant voice would significantly decrease the non-caregiving behavior regardless of the robot form. The results support our hypotheses H2. The results of this study suggest that increasing the baby-like features of different modalities would promote caregiving behavior, while decreasing the non-caregiving behavior. On the other hand, the visual modality did not have effects on the number of different kinds of behaviors as expected. It can be concluded that the behaviors did not increase or decrease linearly, rather a complex effect on those behaviors. Significant differences were found between baby-form and all other forms, yet no other differences were observed on other forms. From this result, we can see the importance of having a limb-like structure on the morphological design for the robot to receive more caregiving behavior and less non-caregiving behavior. From the results, we can also see that the interaction between robot forms and voices was not significant for any kind of behavior,

though either form or voice was significant alone. We believe that this could be because of the mismatch in the form and voice expressions. While the robot forms used in this experiment were rather abstract, the voices used in this experiment were actual recordings. The difference in these stimuli of baby-like modalities may have caused the mismatch. Psychological research on audiovisual integration in emotion has reported that people show visual dominance in emotion recognition, although when the visual information's reliability is low, they recognize emotions using more reliable auditory information (Collignon et al., 2008). Since all the robots used in this study had an abstract baby form, the baby voice's effect was stronger than the visual information. It would be interesting to investigate how baby-like features affect the caregiving and non-caregiving behaviors when more realistic baby forms are used, or if baby-like characteristics were removed from the baby's voice.

Differences in human behavior induced by multimodal baby features

One possible explanation for the lack of interaction between the robot's physical form and its voice is that bimodal cues—those involving both visual and auditory information—were only present in a subset of the actions and speech components that comprised each behavior. To explore this further, we conducted a detailed analysis. As stated in the previous parts, we separated caregiving behaviors into infant directed actions and infant directed speech, and non-caregiving behaviors into non-infant directed actions and non-infant directed speech. This analysis revealed a significant interaction between the robot's form and its voice specifically within the infant directed action condition. Post-hoc comparisons indicated that in the absence of speech, the baby-shaped robot

elicited significantly more caregiving behavior than all other shapes, replicating the main effect found for caregiving behavior. However, when speech was present, significant differences emerged only between the baby-form and round-form robots, and between the square-form and round-form robots; no other comparisons reached significance. These results suggest that the presence of a baby voice compensated for the lack of baby-like visual features in some robot forms (e.g., cross, snow, square), thereby enhancing the likelihood of eliciting caregiving behavior. This complementary effect appeared to rely on whether the robot's form could stimulate participants' imagination through auditory baby cues—a point supported by post-experiment interviews. As a reference, one participant mentioned that the square-formed robot is easily imagined as an infant in a baby crib. In contrast, the round-form robot was often imagined to be a pillow or cushion rather than an infant, which could have led to lower levels of caregiving behavior. These findings imply that even when a robot utilizes baby voice, it may not be enough to trigger caregiving behaviors if its appearance does not afford imaginative associations with an infant. In the meantime, in the analysis of infant directed speech, we did not find such interactions. We found no difference in the form of robots. The only difference we found was the number of speeches depending on whether the robot possesses an infant voice. We analyzed the speech of the participants. The results showed that participants respond to the voice of the robots regardless of what form the robot is. It is rarely observed where a participant actively talked with the robot. Though there have been results on communication robots where a more friendly appearance would increase the communication between the robot and the user (Mathur & Reichling, 2016). However, these results were based on the fact that the robot was highly intelligent and had the ability to socialize. The

robots were expected to response to the speeches of the participants. Therefore, the friendly appearances were able to increase the amount of communication from the participants. In this experiment, participants were interacting with a baby-like robot, which was not expected to make the appropriate responses by talking to it. Since in this experiment, all the participants were younger adults under the age of 50, talking to an infant without the ability to respond may seemed wired to them. The result could appear different in seniors, and we dug deeper into this topic in the later stages of our study.

Effects of Non-Infant Directed Actions and Speech

In contrast to the findings for infant directed actions, no comparable interaction effects were observed in the case of non-infant directed actions. This suggests that the presence of multimodal baby-like features alone is insufficient to suppress negative user behavior. In other words, reducing such behavior may require a simple increase in the degree of baby-like characteristics. Regarding speech, similar to the results found in infant directed speech conditions, the robot's voice emerged as the only significant main effect in the non-infant directed speech condition. An analysis on the non-infant directed speech showed that those speeches were relatively rarely made by participants compared to the infant directed speech and mainly consisted of self-directed comments showing the confusion either about the robot, how to properly interact with it or on the situation. When the robot's voice exhibited baby-like qualities, participants were less likely to engage in such self-talk, likely due to a clearer perception of the robot as an infant. Conversely, the absence of baby-like vocal cues led to more frequent expressions of confusion, indicating a lack of clarity in how the

robot should be approached.

Absence of Gender Differences in Caregiving and Non-Caregiving Behaviors

The study found no significant gender differences in either caregiving or non-caregiving behaviors. Male and female participants demonstrated comparable levels of caregiving behavior under the same experimental conditions. This result remained consistent across subcategories, including both physical actions and verbal expressions. Both male and female participants engaged with the baby-type robot in similar ways, reflecting no substantial variation in their interactions. These findings align with the observations reported by Rutherford and Przednowek (2012), who noted no gender differences in parent-infant interactions. Our results suggest that caregiving behavior toward infant-like robots mirrors human caregiving toward infants and does not necessitate gender-specific design modifications. Although not originally specified in the research questions, the study also explored whether experience with children affects caregiving and non-caregiving behaviors. An additional four-factor ANOVA was conducted, incorporating participant gender, robot voice, robot form, and participants' prior experience with raising or recently playing with children. The analysis revealed no significant main or interaction effects related to caregiving experience, suggesting that such experience does not influence caregiving responses. These findings provide indirect evidence that caregiving behavior is likely an instinctive human response.

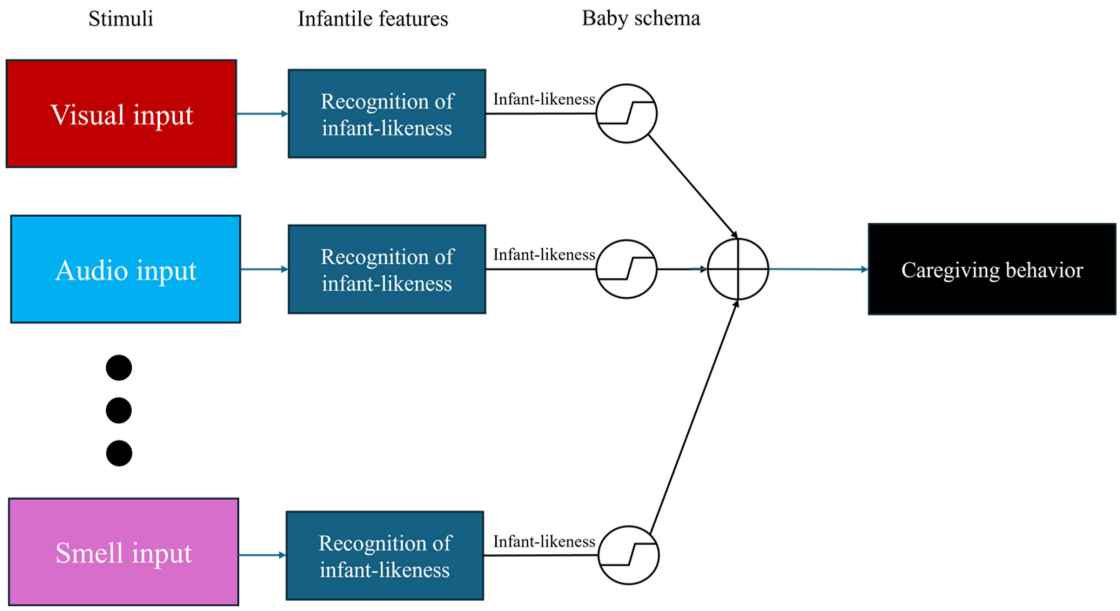


Figure 2.12: Model of triggering caregiving behavior

A Multimodal Baby Schema Model for Eliciting Caregiving Behavior

Drawing on the overall results, we propose a conceptual model describing how multimodal infant-like features elicit caregiving behavior (see Fig. 2.12). This model integrates sensory cues from multiple modalities—visual appearance, vocal characteristics, olfactory signals, and others—that are commonly associated with human infants. When exposed to these stimuli, the human brain evaluates whether the target entity possesses infant-like qualities. This evaluation process is nonlinear and cannot be understood as the mere summation of independent features. Instead, the perception of an entity as ”infant-like” emerges from the integration of multimodal sensory inputs. Once this holistic judgment is formed, caregiving behaviors are activated in response to the perceived infant schema.

Limitation and future work

This study has several limitations. First, the multimodal features were not well matched. The robot had a cartoon-like appearance but used a realistic voice. This mismatch may have confused the participants and reduced the baby schema effect. Second, we studied how the loss of visual features changed the perception of the baby schema, but we did not test the same change in voice. Future work will add cartoon-like voices to balance the audio and visual features. This may provide more insights into how multimodal features can trigger the baby schema. Third, the participants were all younger than 50 and had no mental health problems. Future work should include older adults and compare the results between younger and older generations, as well as examine whether the findings apply to both healthy seniors and seniors with dementia. Fourth, we did not control the emotional state of the robot. The robot could cry, but the participants did not have the chance to respond to it. Emotional states of the robot may change human behavior. In this study, we only contributed to the morphological design of baby-like robots and Hiro-chan to a general extent. We found that the current Hiro-chan design is sufficient and baby-like robots should have illustration of infant limb and infant voice. However, we did not specify in detail how minimized baby-like robot designs can be to achieve enough efficiency. In the future, after choosing a suitable multimodal design, we plan to study how emotional design affects the interaction and discover more details to quantify the boundary of minimal design of baby-like robots. Also, in the behavioral analysis we ran separate ANOVAs on three measures (caregiving, infant-directed action, infant-directed speech). Because caregiving equals action plus speech, the two component measures were reused across tests. To control multiplicity, we treated tests of the same effect across the three measures as one

family ($m = 3$) and applied the Holm–Bonferroni procedure ($\alpha = .05$). Within each family, the ordered p-values $p_{(1)} \leq p_{(2)} \leq p_{(3)}$ were compared to the stepwise thresholds .0167, .0250, and .0500, respectively. After correction, the main effect of voice remained significant for caregiving, action, and speech (all $p < .001$). The main effect of form remained significant for caregiving and action ($p < .001$) but not for speech. For the Form × Voice interaction, the smallest uncorrected p-value was $p = .039$, which failed the first-step threshold (.0167); therefore, none of the three interaction tests is significant after correction. For the non-caregiving behavior, non-infant directed action and non-infant directed speech, the main effect of robot form and the robot voice remained significant after the correction (all $p < .001$. Due to such changes in the results after the correction, the discussion made in this study We built the inferred model of caregiving behavior where multimodal infant features in baby-like robots form a sigmoid-like function to trigger baby schema and all baby schema all together contributes to infant directed actions, and the same stands for infant-directed speech. However, since no interaction between robot form and voice was spotted after the holm-Bonferroni correction, the inferred model can also be changed and become the same model as the non-caregiving behavior, where baby schema effect contributes to non-infant directed action and speech with a NAND gate process. In the future, more data on this subject and the experiments using more different infant features can be conducted to confirm and verify the model. Finally, in this study, we only contributed to the morphological design of baby-like robots and Hiro-chan to a general extent. We found that the current Hiro-chan design is sufficient and baby-like robots should have illustration of infant limb and infant voice. However, we did not specify in detail how minimized baby-like robot designs can be to achieve enough efficiency. In the future, after choosing

a suitable multimodal design, we plan to study how emotional design affects the interaction, and discover more details to quantify the boundary of minimal design of baby-like robots.

CHAPTER 3

MORPHOLOGICAL DESIGN OF BABY-LIKE ROBOTS WITH OLDER ADULTS

3.1 Current study

To find out the relationship between the extent of multimodal infant features in the design of baby-like robots and the number of caregiving behaviors and the baby schema effect, experiment using five different levels of morphological designs on infants and the on and off of recordings of real infants on the robots as the variables, questionnaire and behavioral analysis was conducted on the study group of younger adults under the age of 50. The study established that a threshold of infant features exists where baby schema effect can be triggered when a certain feature exceeds the threshold. The caregiving behavior is related to the addition of all the baby schema effect triggered by each infant feature [10]. This result is valid among younger adults. Though we do not believe caregiving behavior can differ between different age groups since it is intuitive behavior, the recognition on baby-like robots can differ with the difference in cognitive and emotional processing abilities. Studies have shown that older adults have regressed levels of cognitive and emotional processing abilities than younger adults [11, 12]. Such differences can cause the recognition of infant features to be different, thus having different results than younger adults. In light of this possibility, we felt the necessity of conducting the study on older adults as well to see how the cognitive and emotional processing changes the baby schema and caregiving behavior induced by infant features in baby-like robots.

3.2 Experiment

After completing experiment I, we understood some traits and better morphological design of baby-like robots for younger adults. To thoroughly understand the morphological design of baby-like robots, we conducted the same experiment on older adults as well. In the following parts, we will be introducing the purpose, process, results and discussions of the experiment on older adults. This experiment is named experiment II. In experiment II, we utilized the same five forms of robots. Since we have understood the significance of voice in experiment I, the robots used in experiment II were all with voice function on.

3.2.1 Participants

The participants of this experiment were all older adults. We experimented on 38 older adults in total. Within them, 21 were female and 17 were male. The age of the participants varied from 60 to 80, where all of them were healthy adults without dementia and average age is 72.5 (SD=5.08). Since the experiment and the questionnaire, interview after the experiment was conducted in Japanese language, the recruits all have native Japanese level. Other than the age and the gender, there was no excluding process for the participants during recruiting. The participants were recruited from the silver human resources center in Seikacho. The sample size was designed using the same method as in Chapter 2.

3.2.2 Experimental design

In this experiment, the same five forms of robots were used to understand the morphological design. Since we have already proved the influence of infant voice, all the participants interacted with the robots with the voice function on. For the result of the experimental design, we remained the same as in Chapter 2. In order to compare the results with the results in younger adults, we kept the sample size of older adults similar with the number of participants in group V-on in experiment I, while also conducting a reverse ANOVA mentioned in experiment with younger adults. The sample size is then determined to 45.

3.2.3 Hypothesis

In this study, we aim to understand how infant features in the design of baby-like robots affect caregiving behavior in older adults. Since we have already studied that of the younger adults, in this study, we also aim to compare the differences between younger adults and older adults in their opinions and behaviors facilitated by infant features in baby-like robots. From the studies mentioned above, including the study on younger adults, we made the following hypothesis H1: Older adults are more forgiving in the multimodal infant features in baby-like humans than younger adults. In the study on the relationship between multimodal infant features and caregiving/non-caregiving behavior with younger adults, we proposed the hypothesis where robots with more infant features increases the number of caregiving behavior and reduces non-caregiving behavior towards the robots with less infant features, and the two hypotheses were supported by the results of the same study. Since we believe

that older adults may have less cognitive capability for recognizing infant features, and that they have more experience with children, they are likely to be acceptable to baby-like robots with less infant features. H2: Older adults perform more caregiving behavior and less non-caregiving behavior on all kinds of baby-like robots than younger adults. Since older adults generally have more experience with children than younger adults, and older adults have less experience in knowledge about modern technology such as robots, they are more likely to treat the baby-like robots more gently. That is why we made this hypothesis. H3: Male participants show the same adaptions of caregiving and non-caregiving behavior as female participants regarding all the infant features for older adults. In the experiment with younger participants, we did not find and significant difference in impressions and behavior towards baby-like robots among gender, which is consistent with the research by Rutherford & Przednowek, 2012, where male and female show similar adaptions in behavior towards infants. Though we have considered the possibility that with more experience with children, older adults may have a stricter standard towards baby-like robots, results in the younger adults showed that experience with children did not have any impact. We assume the same for older adults. This is why in this study, we presume the same for older adults.

3.2.4 Procedure

In this experiment, we followed the same procedure as in experiment with younger adults. To summarize, we took the consent from the participant, led the participant into the experiment room, and explained the content of the experiment. Then the experimenter took the first robot inside, then left the room.

After approximately one minute, the experimenter came back in, took the robot away and handed over the next robot. The process was repeated until all five forms had been interacted with. After the interaction, the experimenter reentered the room with all five robots lined up on the table in the same order the participant interacted with, and took the questionnaire, interview. After this, the experiment ended. The process was the same for all participants.

3.2.5 Data processing

In this experiment, the data collected were the same as that in experiment with younger adults, which includes the questionnaire results, the interview answers and the video recording. The processing of ranking results was the same as experiment I, converting ranking into scores. The questionnaire and behavioral result analysis was also using ANOVA. Since in this experiment, we only used robots with voice on, we deleted the between-subject factor of group V-on and V-off.Thus, in this experiment, we conduct the three-factor mixed ANOVA with between-subject factors of 1) participant gender; 2) experience with children. The within-subject factor is the robot forms. The classification of caregiving and non-caregiving behaviors were also the same as in experiment I.

3.2.6 Results

In the results part of experiment II, we will be introducing the results on questionnaires and the behavioral analysis.

Baby likeness

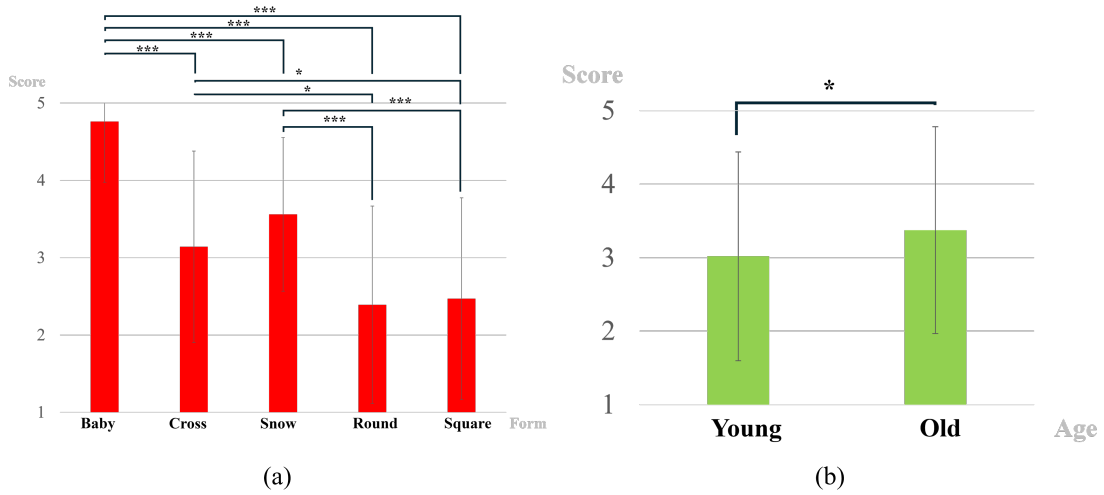


Figure 3.1: (a) Baby likeness scores among all robot forms; (b) baby likeness scores against participant age
$(* : 0.01 < p < 0.05, ** : 0.001 < p < 0.01, * * * : p < 0.001)$

Figure shows 3.1 the results of baby likeness scores in the questionnaire. Significant differences in main effects of robot forms ($F_{(4,248)} = 49.97, p < .001, \eta_p^2 = .45$) and participant age ($F_{(1,62)} = 5.51, p = .022, \eta_p^2 = .08$) are found. Post-hoc comparison with Bonferroni corrections showed that baby-form ($M = 4.79, SD = 0.10$) has significantly more scores than cross-form ($M = 3.01, SD = 0.16, p < .001$), snow-form ($M = 3.48, SD = 0.13, p < .001$), square-form ($M = 2.36, SD = 0.17, p < .001$) and round-form ($M = 2.34, SD = 0.17, p < .001$). Cross-form has significantly better scores than square-form ($p = .042$) and round-form ($p = .042$). Snow form has significantly better scores than square-form (p¡ .001) and round-form ($p < .001$).

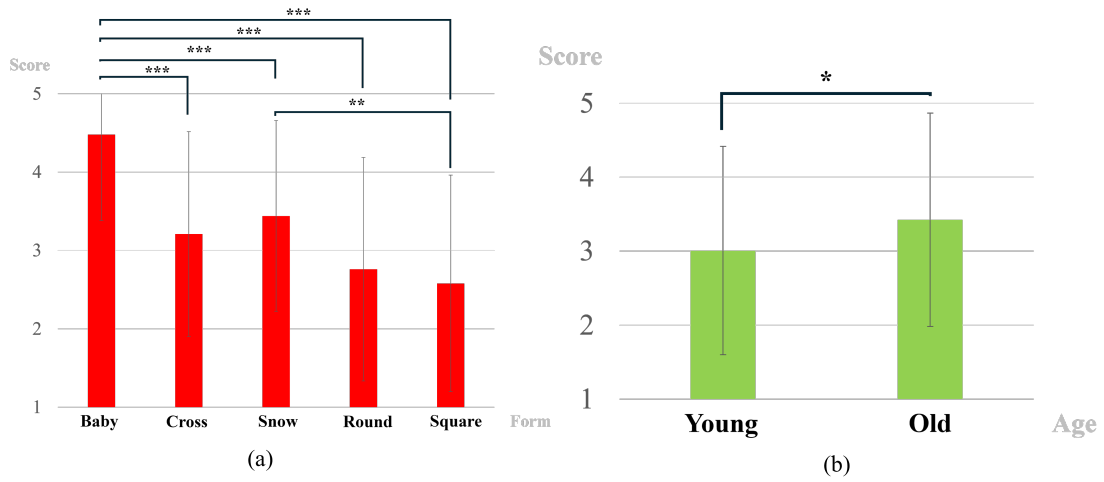


Figure 3.2: (a) Fun to play scores among all robot forms; (b) fun to play scores against participant age
$(* : 0.01 < p < 0.05, ** : 0.001 < p < 0.01, *** : p < 0.001)$

Fun to play with a robot

As can be seen in Figure 3,2, we found main effect in the fun to play scores of robot form ($F_{(4,248)} = 21.37, p < .001$, $\eta_p^2 = .26$) and participant age ($F_{(1,62)} = 5.93, p = .018$, $\eta_p^2 = .09$) significant. Post-hoc comparison with Bonferroni corrections showed a significant better score on baby-form ($M = 4.43, SD = 0.14$) than all other forms including cross-form ($M = 3.09, SD = 0.17, p < .001$), snow-form ($M = 3.33, SD = 0.16, p < .001$), square-form ($M = 2.47, SD = 0.18, p < .001$) and round-form ($M = 2.76, SD = 0.19, p < .001$). Square-form has significantly less scores than round-form ($p = .004$).

Easy to play with a robot

As can be seen in Figure 3.3, we found main effect of robot form ($F_{(4,248)} = 29.28, p < .001$, $\eta_p^2 = .32$) and participant age ($F_{(1,62)} = 4.32, p = .042$, $\eta_p^2 = .07$) significant. Post-hoc comparison with Bonferroni corrections showed a signif-

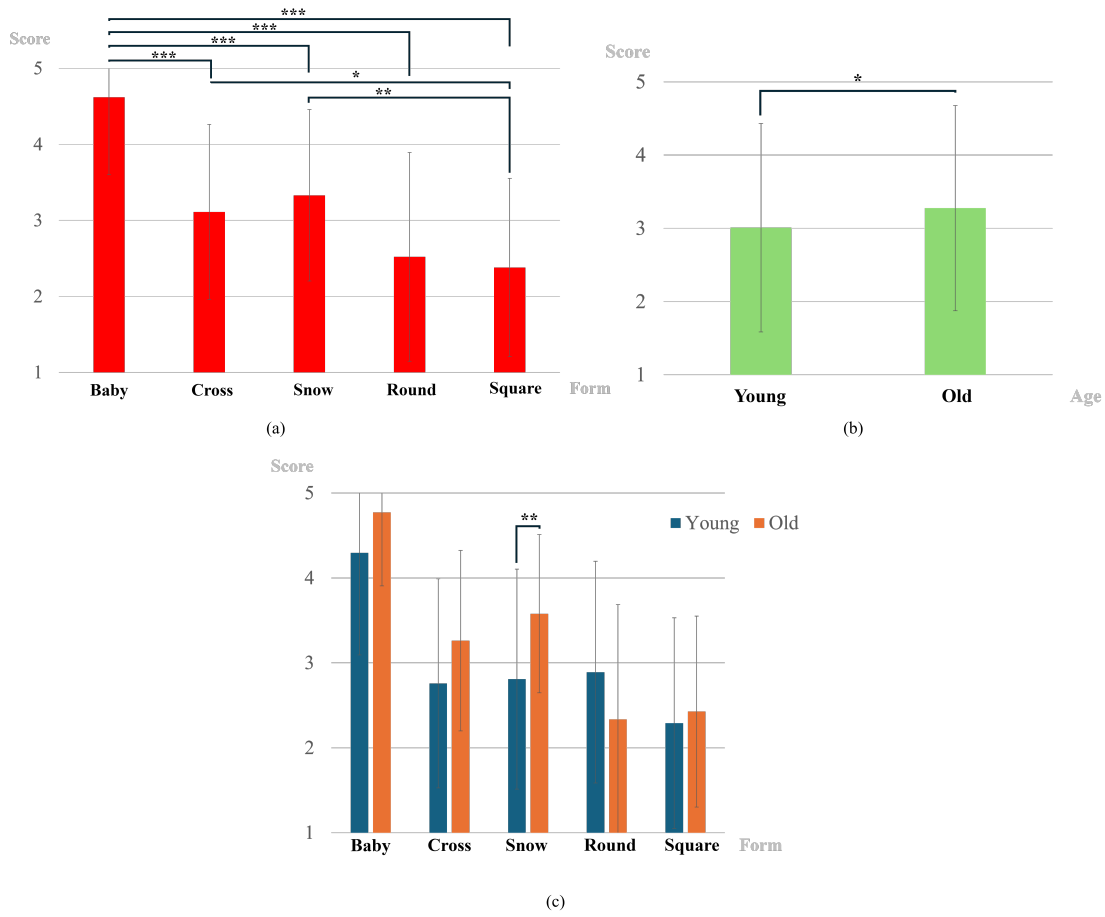


Figure 3.3: (a) Easy to play scores among all robot forms; (b) easy to play scores against participant age; (c) easy to play scores among robot form x participant age
$(* : 0.01 < p < 0.05, ** : 0.001 < p < 0.01, * * * : p < 0.001)$

icant better score on baby-form ($M = 4.534, SD = 0.13$) than all other forms including cross-form ($M = 3.01, SD = 0.15, p < .001$), snow-form ($M = 3.19, SD = 0.15, p < .001$), square-form ($M = 2.36, SD = 0.16, p < .001$) and round-form ($M = 2.61, SD = 0.18, p < .001$). Square-form has significantly less scores than cross-form ($p = .024$) and snow-form ($p = .003$). Significant difference was also observed in the interaction between robot form and participant age with $F_{(4,248)} = 2.69, p = .032$, $\eta_p^2 = .04$. Post-hoc found difference ($p = .010$) between the scores of snow-form of younger ($M = 2.81, SD = 0.24$) and older

($M = 3.58, SD = 0.16$) participants.

Caregiving behavior

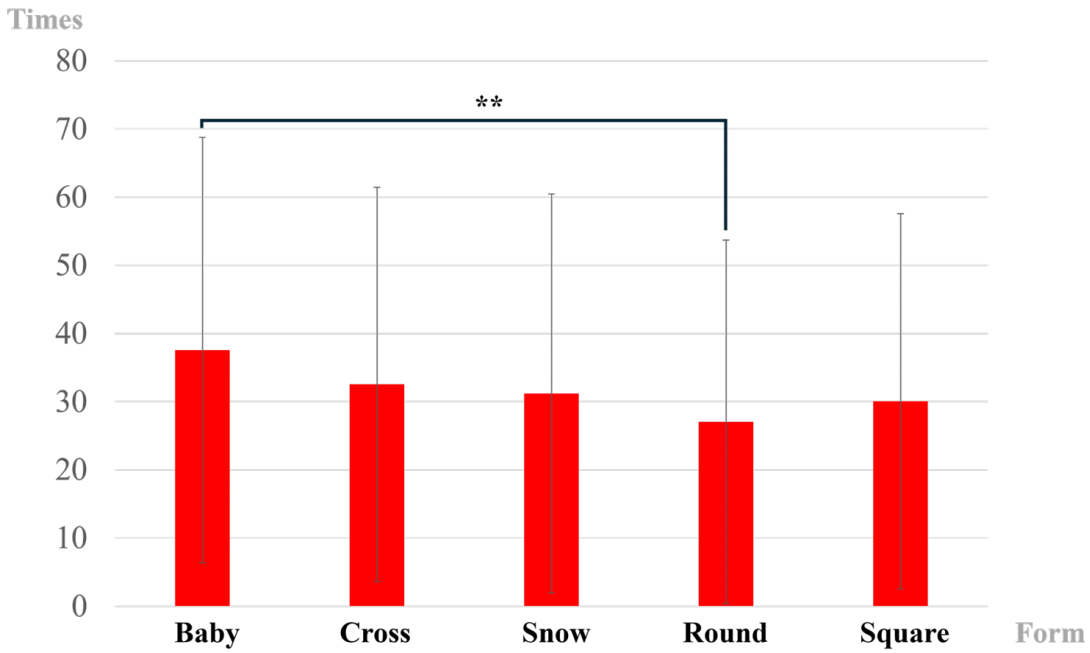


Figure 3.4: Caregiving behavior among all robot forms
($* : 0.01 < p < 0.05, ** : 0.001 < p < 0.01, * * * : p < 0.001$)

Total behaviors Figure 3.4 shows the results of caregiving behaviors. We found the main effect of caregiving behavior with robot form ($F_{(4,236)} = 4.28, p = .002$, $\eta_p^2 = .07$). Post-hoc comparison with Bonferroni corrections showed that caregiving behavior is observed on baby-form ($M = 36.22, SD = 4.01$) significantly more than on round-form ($M = 26.21, SD = 3.39, p = .010$).

Infant directed actions The main effect of age is significant in infant directed action ($F_{(1,59)} = 13.59, p < .001$, $\eta_p^2 = .19$). Older participants ($M = 13.47, SD = 1.91$) significantly performed more infant directed actions than younger partic-

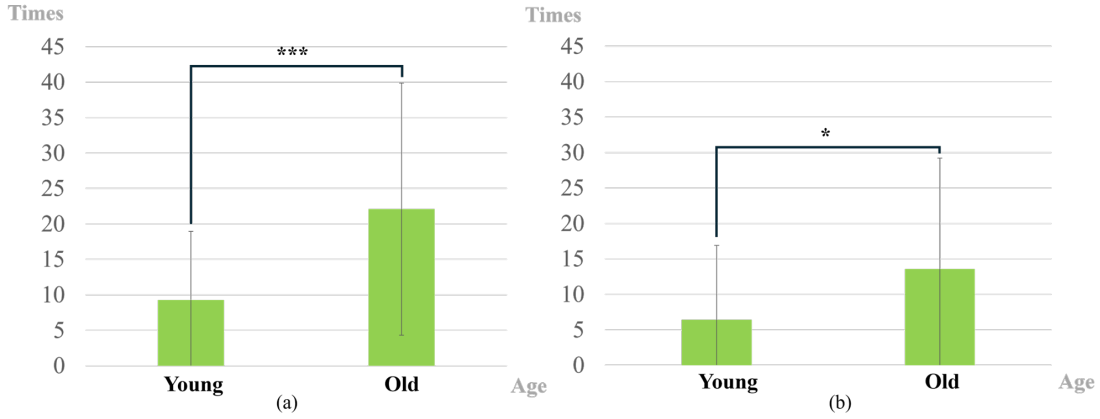


Figure 3.5: (a) Infant directed action among participant age; (b) infant directed speech among participant age
($* : 0.01 < p < 0.05, ** : 0.001 < p < 0.01, * * * : p < 0.001$)

ipants ($M = 6.55, SD = 2.35$). We found no significant difference in robot form, age or gender (Fig. 3.5 (a)).

Infant directed speech The main effect of age ($F_{(1,59)} = 5.22, p = .026, \eta_p^2 = .08$) is significant in infant directed speech. Older participants ($M = 22.07, SD = 2.22$) significantly used more infant directed speech than younger participants ($M = 9.75, SD = 2.72$). We found no significant difference in robot form, age or gender (Fig. 3.6 (b)).

Non-caregiving behavior

Total behaviors Non-caregiving behaviors showed main effect with robot form ($F_{(4,236)} = 6.05, p < .001, \eta_p^2 = .09$). Post-hoc comparison with Bonferroni corrections showed that caregiving behavior is observed on baby-form ($M = 4.29, SD = 0.71$) less than all other robot forms. Participants performed less non-caregiving behaviors to baby-form than cross-form ($M = 7.60, SD =$

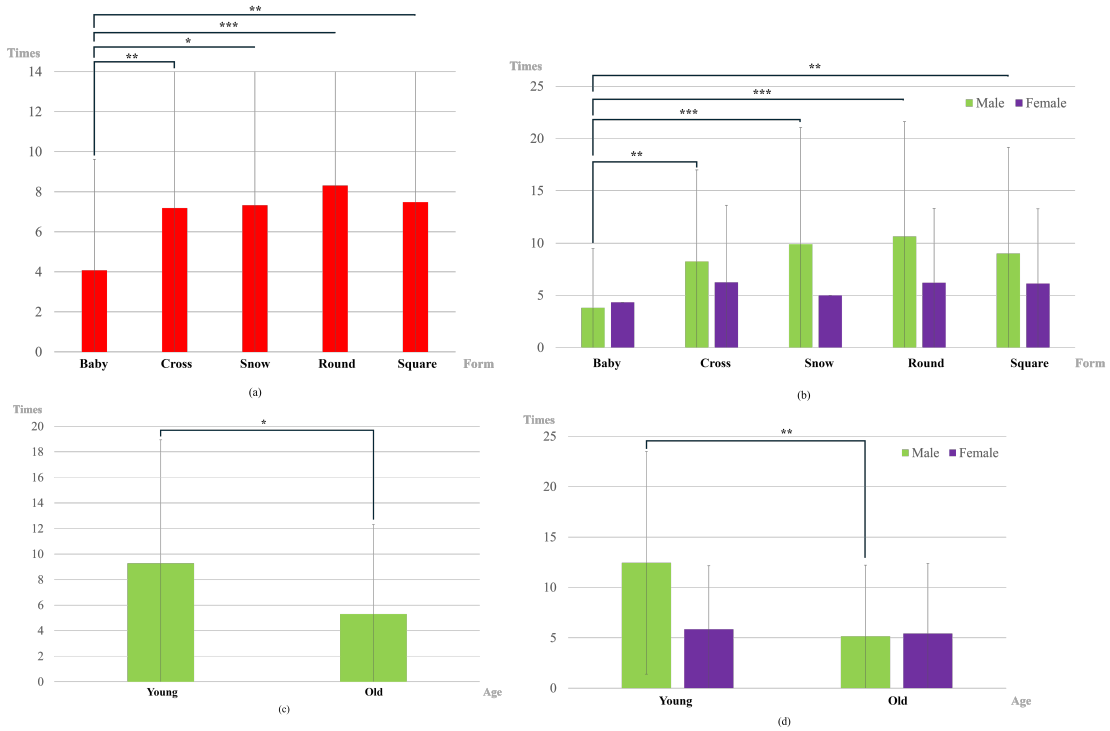


Figure 3.6: (a) Non-Caregiving behavior among all robot forms; (b) non-Caregiving behavior among participant gender with all robot forms; (c) non-Caregiving behavior among participant age; (d) non-Caregiving behavior among participant age with participant gender
$(* : 0.01 < p < 0.05, ** : 0.001 < p < 0.01, *** : p < 0.001)$

$1.00, p = .001$), snow-form ($M = 7.55, SD = 1.16, p = .014$), round-form ($M = 8.91, SD = 1.05, p < .001$), and square-form ($M = 7.74, SD = 1.11, p = .005$). We also found main effect on the interaction between robot form and participants' gender ($F_{(4,236)} = 3.03, p = .018, \eta_p^2 = .05$). Male participants showed significant different, where female participants did not. In male participants, non-caregiving behavior is observed most in baby-form ($M = 4.04, SD = 1.02$), showing significant different against cross-form ($M = 8.73, SD = 1.42, p = .001$), snow-form ($M = 10.31, SD = 1.65, p < .001$), round-form ($M = 11.49, SD = 1.50, p < .001$) and square-form ($M = 9.43, SD = 1.58, p = .002$). We also found significant main effect of participant age different ($F_{(1,59)} = 5.83, p = .019, \eta_p^2 = .09$). The interaction between participant age and participant gender is

also significant ($F_{(1,59)} = 4.62, p = .036, \eta_p^2 = .07$). Post-hoc comparison with Bonferroni corrections showed that in male participants, younger participants ($M = 12.45, SD = 11.05$) significantly performed more non-caregiving behavior than older participants ($M = 5.15, SD = 7.07, p = .002$).

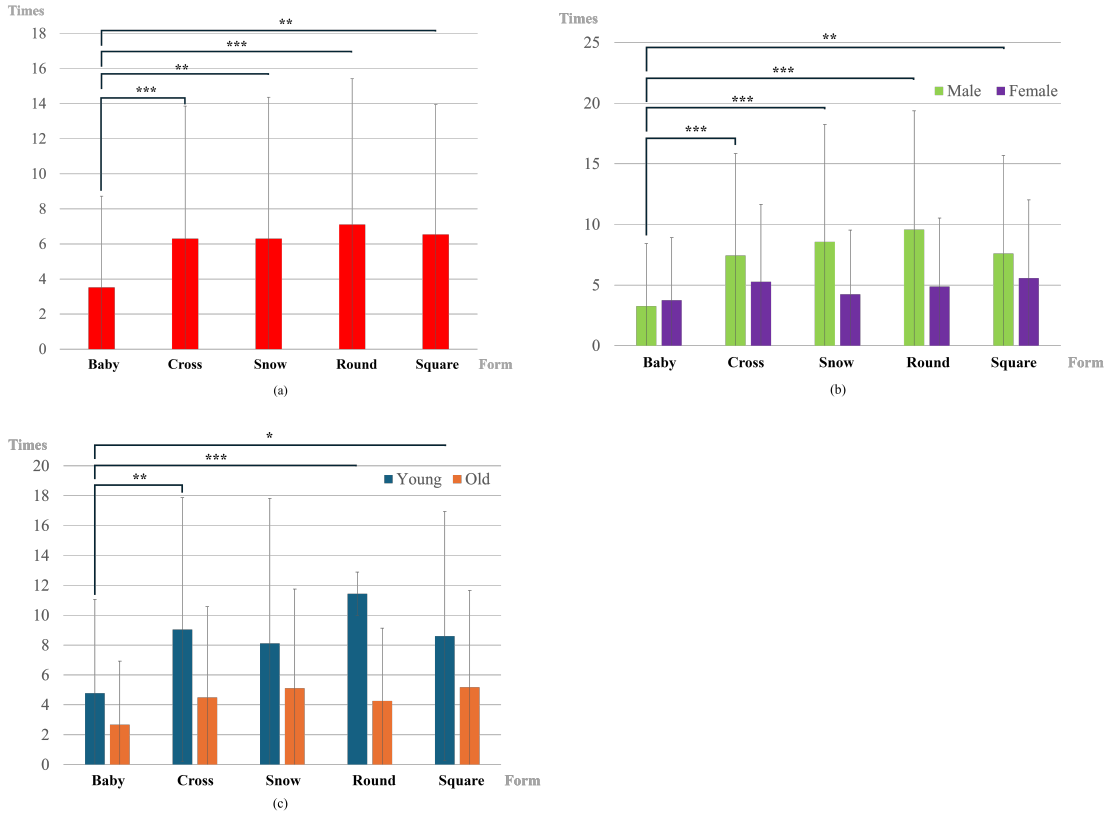


Figure 3.7: (a) Non-infant directed action among participant gender with all robot forms; (b) Non-infant directed action among participant gender with all robot forms; (c) Non-infant directed action among participant age with all robot forms
($* : 0.01 < p < 0.05, ** : 0.001 < p < 0.01, *** : p < 0.001$)

Non-infant directed actions The main effect of robot form ($F_{(4,236)} = 7.14, p < .001, \eta_p^2 = .11$) is significant in non-infant directed action. Post-hoc comparison with Bonferroni corrections showed that non-infant directed action is performed significantly less on baby-form ($M = 3.69, SD = 0.67$) than all other forms: cross-form ($M = 6.72, SD = 0.94, p < .001$), snow-form ($M = 6.56, SD =$

$0.99, p = .008$), round-form ($M = 7.76, SD = 0.90, p < .001$), square-form ($M = 6.82, SD = 0.92, p = .002$). There are also significant differences in interaction between participant gender and robot form ($F_{(4,236)} = 3.86, p = .005$, $\eta_p^2 = .06$), and interaction between participant age and robot form ($F_{(4,236)} = 2.67, p = .033$, $\eta_p^2 = .04$). Post-hoc comparison with Bonferroni correction showed that there is no significant difference in robot forms among female participants. In male participants, non-infant directed action is performed less on baby-form ($M = 3.50, SD = 0.94$) significantly than all other forms: cross-form ($M = 7.89, SD = 1.33, p < .001$), snow-form ($M = 9.00, SD = 1.41, p < .001$), round-form ($M = 10.36, SD = 1.28, p < .001$), square-form ($M = 8.04, SD = 1.32, p = .001$). In the interaction between participant age and robot form, the Post-hoc comparison showed that there is no significant difference among older participants. Among younger participants, non-infant directed action is performed on baby-form ($M = 4.78, SD = 1.04$) significantly less than cross ($M = 8.94, SD = 1.45, p < .002$), round ($M = 11.24, SD = 1.39, p < .001$) and square ($M = 8.49, SD = 1.43, p < .031$) forms.

Non-infant directed speech We found no main effect that is significant in non-infant directed speech.

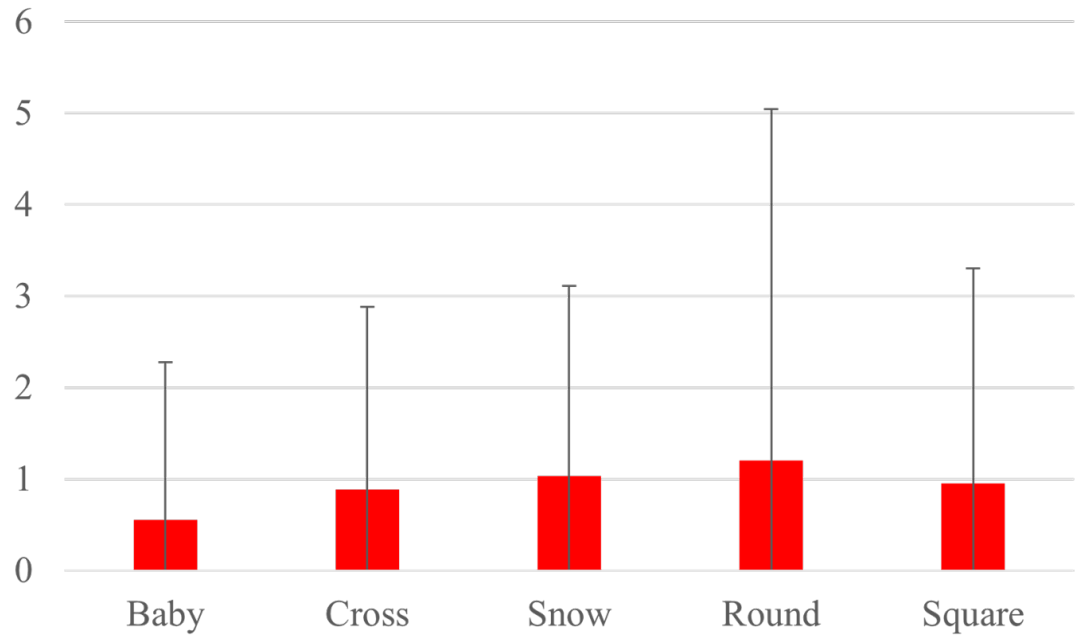


Figure 3.8: Non-infant directed speech in all participants among all robot forms ($* : 0.01 < p < 0.05, ** : 0.001 < p < 0.01, * * * : p < 0.001$)

3.2.7 Discussion

3.2.8 Implications

We changed the shape of a baby-like robot to observe how the morphological design affects the impression and the behavior of older adults. Unlike the previous study on younger adults, we limited the variation to shape only. The results of the questionnaire showed that older adults have similar attitudes towards baby-like robots, where they like the baby-form the most, over all other forms, and that they like snow-form the second, over square-form and round-form. In the meantime, older adults have significantly better attitude towards baby-like robots than younger adults for all the questions. There was no interaction between the robot forms and age except for the snow-form in the easy

to play scores, suggesting that no matter the forms, older adults gave better scores. These results support our H1 where older adults have same standards but better impressions on the baby-like robots. In the meantime, we found that with realistic infant voice, shapes of baby-like robots do not really have much impact on caregiving behavior unless the shape is extremely not baby-like, such as round-form. There is no difference in the tendencies of caregiving behavior between younger and older adults except that older adults generally perform more caregiving behavior than younger adults. The difference in behavior of younger and older adults lies in the non-caregiving behavior. In non-caregiving behavior, the effects of robot shapes differ more with baby-form better than all other forms. Significant difference in the interaction between robot form and participant age suggests that the form difference is more relevant to the non-caregiving behavior for younger adults than older adults. Older adults do not respond negatively to different shapes while younger ones do. After breaking down the non-caregiving behavior into non-infant directed actions and non-infant directed speech, we found that the difference mainly occurred in the actions. We found no difference in non-infant directed speech while difference in shape, interactions between shape and gender, shape and age were found. Difference in the effects of multimodal infant features on baby-like robots In the previous paper, we built an inferred model of caregiving behavior, as well as non-caregiving behavior in younger adults. In this study, we compared the results of younger adults and the results of older adults. For caregiving behavior, we believe that the conclusion of needing two infant features is also valid for older adults since we did not find significant difference in the interactions between robot form and participant gender, suggesting no difference in the effect of infant features on baby-like robots on caregiving behaviors for people of dif-

ferent ages. Though we observed a significantly greater number of caregiving behaviors in older adults, the effect of different infant features remains the same for older adults. In the meantime, for non-caregiving behavior, differences in reactions to different shaped baby-like robots were observed. Other than the fact that older adults perform significantly less non-caregiving behavior on robots, they do not show any prejudice for robots with different shapes. While younger adults performed significantly less non-infant directed actions on baby forms than cross, round and square forms, older adults did not show any differences in the experiment. These results suggest that for older adults, the triggering mechanism for caregiving behavior is the same as with the younger adults, while for non-caregiving behavior, older adults show no adaptation in behavior with different robot forms. Older adults are less strict with baby-like robots than younger adults.

Difference in the effects of multimodal infant features on baby-like robots

In the previous paper, we found out the relationship between infant features, baby schema and caregiving behavior. In this study, we compared the results of younger adults and the results of older adults. For caregiving behavior, we believe that the relationship is also valid for older adults since we did not find significant difference in the interactions between robot form and participant gender, suggesting no difference in the effect of infant features on baby-like robots on caregiving behaviors for people of different ages. Though we observed a significantly greater number of caregiving behaviors in older adults, the effect of different infant features remains the same for older adults. In the meantime, for non-caregiving behavior, differences in reactions to different shaped baby-like

robots were observed. Other than the fact that older adults perform significantly less non-caregiving behavior on robots, they do not show any prejudice for robots with different shapes. While younger adults performed significantly less non-infant directed actions on baby forms than cross, round and square forms, older adults did not show any differences in the experiment. These results suggest that for older adults, same standards as younger adults is necessary to trigger caregiving behavior, while the standards to prevent non-caregiving behavior are lower than younger adults. Older adults are less strict with baby-like robots than younger adults.

Gender difference in non-caregiving behaviors

In the previous study, we found no difference in gender for both caregiving and non-caregiving behavior for younger adults. After adding the data for older adults, we test the results on a different gender group. In caregiving behavior, no difference in gender was observed, making the results constant with the study with younger adults. In non-caregiving behavior, we found significant differences in gender and the interactions between gender and shape, as well as in gender and age. These results further expanded the discussion we had in the previous study. Previously, we had a discussion that men and women show no difference in treating baby-like robots, just like they do not show difference treating real babies. However, after considering the results from this study, we believe that for caregiving behavior, men and women do not show any difference. For non-caregiving behavior, women show less prejudice towards shapes of robots than men do. Women perform similar number of non-caregiving behavior on all robots with different shapes as long as infant voice exists. Though

the conclusion on the gender difference changed to some extent, we believe this is still consistent with the study on real infants (Rutherford & Przednowek, 2012). In the mentioned study, they researched the adaption of actions towards infants of men and women. This only consists of elements regarding caregiving behavior, such as repetitiveness, rate and punctuation. There have been no studies on the non-caregiving behavior on babies, and people do not tend to perform such behavior on real babies. We believe that this kind of behavior including violent behaviors and deserting behaviors only occurs for robots, showing that the robot is treated as a machine instead of a baby. Since in this study, we found that women do not show different attitudes towards different kind of robots and men do, this was due to the difference in experience raising children. However, after analyzing the results based on the experience raising children and experience playing with children recently, we found no difference in either of these elements. This further confirmed our conclusion where caregiving behavior is a natural response instead of a trained one.

CHAPTER 4

FUNCTIONAL DESIGN OF HIRO-CHAN

After determining the better morphological design in baby-like robots, concluding a system diagram to illustrate the triggering factors of baby schema effect, which provides a guidance to the minimal design baseline to the morphological design of baby-like robots, we continue to study the functional design of baby-like robots. Other than the morphological design, we believe that the functional design directly influences the impression of the users, as well as the therapeutic effect of baby-like robots. We would like to empirically explore how to utilize the functions of baby-like robots to provide better service for older adults in caring facilities. In this chapter, we conducted an experiment using videos of Hiro-chan to determine the effect of emotions and combinations of Hiro-chan on older adults. We found that the use of two Hiro-chans can enhance the effect of emotions during the interaction, and different emotions are good at achieving different effects for older adults.

4.1 Background

4.1.1 Effect infant emotions on adults

Since there have been many observations on the behavioral change of adults when infants cry or laugh, many studies on the effects of emotions and voices of infants on adults were conducted. Those studies have observed both neurological and hormonal changes in adults along with the emotional expressions of infants, affecting adults in both positive ways and negative ways. Also, adults

respond to different emotions of infants in different ways. For example, when children smile or make an intimate touch with the adults, the adults' rewarding areas in the brain, including ventral tegmental area (VTA), nucleus accumbens and medical OFC activates, increasing satisfaction. Also, for hormones, oxytocin and dopamine increases, increasing the sensation of satisfaction and pleasure, strengthening the emotional bonds between the children and the adults. When children cry, the adults would experience activation of primary and secondary auditory cortex, thalamocingulate circuit, dorsal anterior insula and pre-SMA/SMA & basal ganglia, along with the increase of cortisol and oxytocin. These changes induce the high sensitivity to the children's behaviors, alertness in adults, sympathy and impulse of physical touch with the children. These reactions would help the adults naturally attend to children's needs and safety in time. On the other hand, long term exposure to those neuron activities and hormone increases would lead to aggression, avoidance in care, anxiety and depression in adults as well. Also, studies found female adults experience increase in dopamine when the children cry, making the caregiving behavior more pleasant for them. Rilling, 2013; Swain et al., 2014; Witteman et al., 2019). Other effects can also be found related to the infant, including interactions with the infants and the absence of the infants. As we can conclude from those studies, different emotions and behaviors in infants can lead to drastic differences in the emotional changes in adults in a natural way.

4.1.2 Emotional expressions in baby-like robots

The baby-like robot Babyloid conveys its emotion in two ways, which includes the LED facial expression and the recorded infant voice. The emotions Babyloid

can express includes Lucidity, Sleepiness, Delight, Pain, Sorrow and Shyness. Within those, the emotions that are negative are Pain and Sorrow. While the facial expressions for the two emotions are different, they both use crying to convey the emotions. On the other hand, in the case of Hiro-chan, which is a baby-like robot that is faceless, the sole method to express emotions is via the recorded infant voices. In general, the recorded voices can be classified into three emotions, which are Laughing, Bubbling and Crying. Though different in classification of emotions and methods to convey them, both robots use crying as a way of expressing negative emotions. What is more, it is worth mentioning that crying is not a single emotion, but a way to express negative emotions. Babyloid uses crying to convey two different emotions, while the crying of actual infants can convey more. The message can vary from needing physical care to craving emotional bonds. Meanwhile, with the lack of clear facial expressions, the negative emotions in Hiro-chan rely more on the users' imagination on the crying voice alone. This makes the study of how emotions can be carried out by voice in baby-like robots more important.

4.1.3 Voice functions of baby-like robots

The voice functions of the current baby-like robots usually consist of the recordings of real infants. As mentioned in Chapter 3, the voice real-life infant voices can be a deciding factor for a successful morphological design and have a large impact on the impression of the user as well as the therapeutic effects of the baby-like robots. In general, those real-life infant voices can be classified into three categories. Laughing, bubbling and crying. We believe that those emotions in voices could also have the neurological and hormonal effects on adults

as real infants. Though there have not been any studies measuring neuron activities and hormone changes when interacting with a baby-like robot, similar effects have been discovered during the experiments with baby-like robots. In an experiment where the experimenter used Babyloid to observe how people would react when it cries, complex reactions appeared. The participants felt the urge to both "help the robot" and "anxiety due to the crying robot" . In an experiment using Hiro-chan in a caring facility, robots with crying emotion included experienced a 50% impression decrease after one month of usage, when Hiro-chan robot without crying functions stayed the same level of impression after the same period. Reports of "refusal of care" and "on edged participants" appeared with crying Hiro-chan . As stated above where emotions of infants largely influence the emotions and behaviors of adults, utilizing this trait would make a great difference in the interaction with adults. We believe this trait can also be effective if used well in baby-like robots.

4.1.4 Passive social in socially assistive robots

Other than simply using the voice function as a method of emotional expression to attract and interact with older adults, a conventional method that has been proven effective is the passive social. Passive social refers to the utilization of the verbal or non-verbal interactions between two or more robots for the purpose of providing information to the user. In the interaction, the user would not be directly interacting with the robots but receiving the information the robots provide by watching the interaction between the robots . It is proven that using passive social situations in robots can better broadcast information than using only one robot . This method is already widely used in situations that require

receiving information such as commercials . Other than information broadcasting, the expression of emotions and story plot can also be intensified by using non-verbal interactions in passive social situations . There have been examples of the usage of several Hiro-chans in the caring facility for the older adults passively observed the expression of emotions by multiple robots . With the evidence above, we believe it may be effective to use the passive social situation on baby-like robots to convey emotions of baby-like robots, with two baby-like robots interacting with each other with a participant observing.

4.2 Current study

In this study, we investigate how emotional expressions affect older people and how passive social situation can help with the emotional convey using baby-like robots. It is difficult to formulate a clear hypothesis because of the lack of previous studies on the impact of robot emotional expressions on older users. Therefore, based on the findings of the earlier studies, we examine the following two predictions.

4.2.1 Predictions

Since in the previous studies, attitudes towards a baby-like robot drop when it shows crying expressions, we predict that:
P1: Negative emotional expressions in a baby-like robot will be less attractive to older people than others.
Previous studies about passive-social situations show that the interaction be-

tween multiple robots affects human observers. Therefore, we can expect that the two baby-like robots' combination of emotional expressions holds the same nature. Since previous studies show a strong impact of negative emotions on people, negative emotion may inhibit the effect of other emotions. So, we predict that:

P2: Negative emotional expression in a baby-like robot inhibits the effect of the other emotional expression in passive social situations.

When considering the combination of emotional expressions in passive social situations, we found the need to research the timing in turn-taking of emotional expressions that might influence the result. There have been studies on the time in turn-taking or pauses in human-human and human-robot interactions. Studies show that in conversations between adults, turn-taking usually takes around 200ms [19]. In human-robot interaction, robots are expected to have latencies. The best timing of reaction is around 1s to 2s [20]. A reaction time under one second is considered a lack of sincerity and understanding, whereas a reaction time over two seconds is considered too slow. Depending on the timing of emotional utterances between the two baby-like robots, they may give the impression that they are interacting with each other. Since infants do not apply the social rules of either humans or robots, we wish to find out how to plan the time interval to make the baby-like robots feel like they are interacting with each other without making the robot feel unnatural. Therefore, we address the following questions:

RQ1: Does the timing in turn-taking of emotional expression change the impression of the older people on the baby-like robots?

There has been research on social robots where two robots talking together increases the sense of cuteness. In this experiment, since we are experimenting with one and two Hiro-chan robots, we would like to confirm if the same con-

clusion applies to baby-like robots. On that basis, we also want to confirm other scales besides cuteness. Therefore, we also address the following question:
RQ2: Does using two Hiro-chan robots change the impression of older people on the baby-like robots?
We also investigate the influence of gender on the impression of the robot because a previous study shows gender differences in brain activities when listening to infant crying sounds.

4.3 Experiment I

We experimented to verify our first prediction: negative emotional expressions in baby-like robots will be less attractive to older people than other expressions.

4.3.1 Participants

The participants are older people recruited from a human resource center for the elderly. There are 24 participants (11 male and 13 female). The participants are all native Japanese as the experiment and questionnaire are conducted in the Japanese language. The age of the participants ranges from 61 to 83. The average age is 72 (SD=5.83). The participants are in a mentally healthy condition. The participants understood and agreed to the experiment's procedure and the data collected beforehand.



Figure 4.1: Still image presented in the used video in Experiment I

4.3.2 Experimental setup

As mentioned above, interaction using Hiro-chan without facial expression can still give the participants imagination for the interaction. Using voice can sufficiently make the participants understand the emotion of the robot, and can avoid the strangeness when the facial expression does not match the perceived emotion. In this study, we focus on emotional expression through voice. There-

fore, we used Hiro-chan, a faceless baby-like robot, in this experiment. We also used a video created with still images of the robot and recorded audio to control the stimulus presentation instead of real Hiro-chan robots. There are two main reasons. First, we want the participants to focus more on the difference between the expressions of Hiro-chan. In a pilot study, it is observed that some participants are driven to interact with the robot verbally or physically, focusing less on other factors. Or they would be completely focused on the Hiro-chan robot itself. In this study, we want the participants to focus on the interaction between the robots and how the impression changes when the interaction changes. Although this is an interesting phenomenon to be studied in the future, we believe using the video can help the participants focus more on their impressions, making the experiment result more valid. The second reason is that we want to make sure that the stimuli are safe for the participants. In the previous experiment, some participants were disturbed by the crying sound of Hiro-chan. Since in this experiment, the participants will be exposed to the crying sound of Hiro-chan more, we decided to use the weakened stimuli–video recording, to make sure there is no bad influence before we proceed to experiment using real Hiro-chan robots. Overall, we believe using video to experiment is more fitting in this study as a first step. Also, since the voice function of Hiro-chan consists of real infant voices, in order to make sure the participants have the impression that they are interacting with a baby-like robot, we used the video presentation of Hiro-chan instead of using only audio in the experiment.

The video consisted of a picture of Hiro-chan sitting on the table (Figure 4.1). In the video, a single Hiro-chan is performing its voice function. The recorded infant voice is presented in the video. We prepared three different videos, each of which presents different emotional expressions: crying, laughing, and babbling

sounds (normal), which denote meaningless articulations such as ”ma ma ma.” Each voice sound was presented 200ms after the onset of the video. The audio files are selected from the audio library of Hiro-chan, which are all recorded voices from actual infants. To make sure of the clarity of convey for each emotion, and to make sure the participants receive the same number of stimuli for each kind of emotion, during the selection, we used only the audio files with a reasonable length, and the ones with clear emotions for the participants to understand. The timeline is shown in Figure 3.2. The same kind of emotional voice was repeated 10 times during the video, with 11 audio files in total.

4.3.3 Measurement



Figure 4.2: The timeline of the video of Experiment I

We used some questions used in a study that investigated the impression towards Babyloid before and after interaction and other survey studies on interactive robots to evaluate robot attractiveness. Table 4.1 depicts the list of questions. The first three questions were taken from the study on the impression towards Babyloid mentioned above, and the rest were from studies on other interactive robots. Within those, Conversation refers to the verbal willingness to interact with the robots, and Closeness refers to the willingness to get close to the robots and the willingness to protect the robots. All the meanings of the scales were explained to the participants and the participants understood the

full meaning of them clearly. Though it is generally assumed that other factors such as the experience of raising children could also affect the results. However, in a previous study on the appearance of baby-like robots, the result showed there was no difference in the impression of participants on the child raising experience. This is why we did not add the question on experience raising children in this study. The part measuring intimacy and trust towards the robot is taken into the questionnaire in this study. The participants were asked to answer the questions with a rating from 1 to 7. Rating 1 represents 'Not feeling so at all,' rating 4 represents 'Cannot say either way,' and rating 7 represents 'Strongly feel so.' Though it is generally assumed that other factors such as experience of raising children could also affect the results. However, in a previous study on the appearance of baby-like robots, the result showed there was no difference in the impression of participants on the child raising experience. This is why we did not add the question on experience raising children in this study.

Table 4.1: Questionnaire used in Experiment I (Questionnaire 1).

Question	**Scale**
1. I think the robot is cute.	Cuteness
2. I think the robot is fun to be with.	Fun
3. I would like to interact with the robot.	Interaction
4. I would like to get close to the robot.	Closeness
5. I would like to converse with the robot.	Conversation

4.3.4 Procedure

The procedure was as follows. After a participant arrived at our institute, an experimenter explained the experiment's procedure and the participant gave

written informed consent to join this study. Then, the participant was led into the experiment room with a laptop computer to watch videos and answer questionnaires. After the experimenter instructed how to operate the computer and answered a questionnaire, he left the room. The door to the room was left open, and the experimenter waited near the room. The participant could ask the experimenter questions whenever he/she had trouble operating the computer or answering the questionnaire. After the participant watched a video and completed the questionnaire, the next set of video and questionnaire was presented. The order of the videos was counterbalanced among participants. After watching three videos and completing the corresponding questionnaire, the participant called the experimenter. The experimenter confirmed the completion of the questionnaire, and then the experiment was finished.

4.3.5 Results

Table 4.2: Result of Experiment

Scale	**Laugh**		**Normal**		**Cry**	
	M	**SD**	**M**	**SD**	**M**	**SD**
Cuteness	4.96	1.85	5.21	1.82	4.50	1.89
Fun	4.83	1.83	5.17	1.58	4.04	2.07
Interaction	4.96	1.92	5.00	1.75	4.50	2.04
Closeness	4.96	1.70	4.92	1.58	4.38	2.00
Conversation	4.96	1.78	5.21	1.69	4.25	2.03

Table 4.2 shows the meaning and standard deviation of each emotional expression in each question. We conducted two-way mixed-design ANOVA with within-subject factors of emotional expression (crying, laughing, and normal) and between-subject factors of gender on the rating of the questions. As shown in Figure 4.3, there were significant main effects of emotional expression in

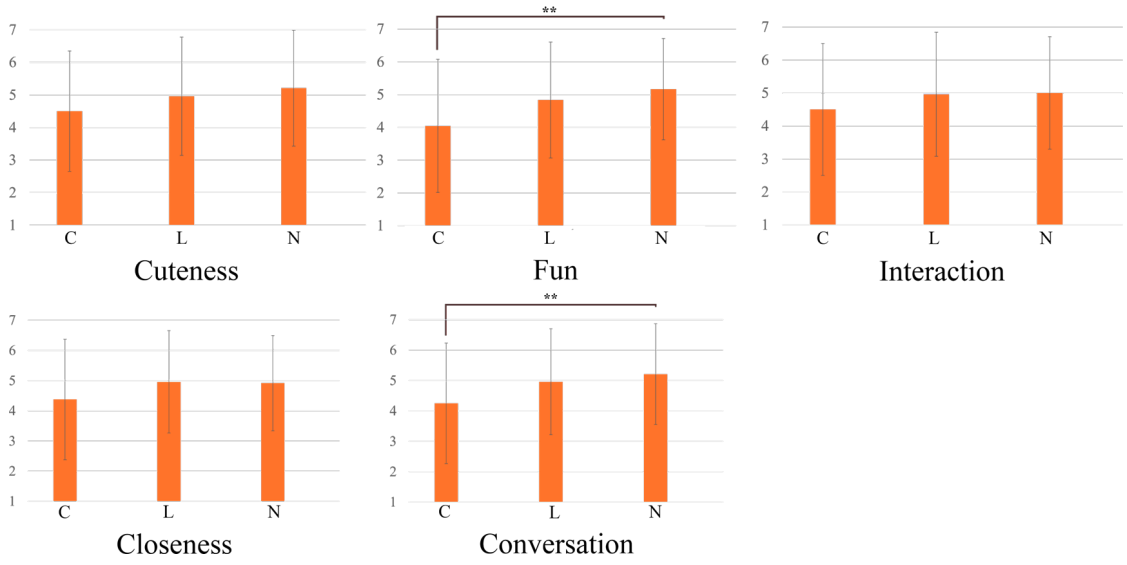


Figure 4.3: The effect of emotional expressions on each scale
($** : p < .01$)

the questions: Conversation ($F_{(2,44)} = 6.40, p = 0.004, \eta_p^2 = 0.23$) and Fun ($F_{(2,44)} = 8.38, p = 0.001, \eta_p^2 = 0.28$). There were no significant differences in gender or the interaction between gender and other elements. Post-hoc comparisons with Bonferroni corrections showed that crying sounds significantly decrease participants' willingness to converse with the robot and feelings of fun than normal sounds. There were no significant effects on the other questions.

4.3.6 Discussion of experiment I

The result of the experiment partially supported our prediction P1. In the first experiment, we found that crying sounds significantly decrease the willingness and fun of older people to interact with robots, compared with normal sounds. However, the effect does not appear to be strong since there were no significant differences between crying and other emotions for most questions. Interestingly, we did not find any significant difference between normal and laughing

for all scales, which indicates neutral, and the laughing of a baby-like robot are the same at attracting older people. Through the experiment, we found partial support for the result that negative attitudes increase when Hiro-chan cries in the nursing home. The result shows that older people like quieter, calmer reactions as well as happy, laughing reactions from baby-like robots. The studies of human infants show that adults react to crying infants in different ways. Some react by performing caregiving behavior while some get annoyed, depressed and even have an urge to harm the infant. We believe that for baby-like robots, users feel less responsibility for caring. Instead of triggering caregiving behavior, the crying of baby-like robots only irritates the user, thus increasing the negative attitudes. The effect of gender is also investigated. The result showed that the gender of older people does not affect the attractiveness of baby-like robots. This result follows the previous findings of males sharing the same attitude and behavior adaption towards infants as females.

The first experiment showed that robot crying negatively affects older people. However, a new question arises in our mind: Will this effect be inhibited by a different emotion, or does this effect inhibit the effect on other emotions? If more crying robots exist, would the negative feelings worsen for older people? In a previous study, there existed several robots on the same floor in a nursing home. In such situations, some robots are crying, others are laughing, the others are saying babbling words. Do crying robots inhibit the effect of other emotional expressions? How does the combination of emotional expressions affect older people? We conducted the second experiment to answer these questions.



(a) Image in pause



(b) Image when left robot is making utter emotional voices



(c) Image when right robot is making utter emotional voices

Figure 4.4: Still image presented in the video used in Experiment II

4.4 Experiment II

We conducted a second experiment to verify our prediction, P2: Negative emotional expression in a baby-like robot inhibits the effect of the other emotional expression in passive social situation., and RQ1: Does the timing in turn-taking

of emotional expression change the impression of the older people on the baby-like robots?

4.4.1 Participants

The same group of participants in Experiment I continued to participate in Experiment II. One member (female, aged 65) could not finish Experiment II due to poor health conditions. During the experiment, the experimenter observed that the participant appeared unfit and was losing focus on the task. The experimenter called off the experiment and sent the participant back to rest. As a result, there were 11 male and 12 female participants in Experiment II. The average age is 72 (*SD*= 5.76).

4.4.2 Experimental setup

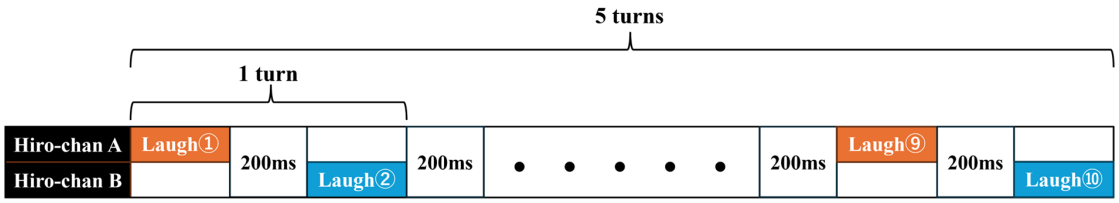


Figure 4.5: The timeline of the video of Experiment II (200ms)

For the same reason as Experiment I, we used the video made with a still image of Hiro-chan and recorded audio. Since the number of audio files is limited in experiment I, the number of time gaps for each emotion is also the same in this experiment. Experiment II is also conducted with the assistance of the researchers on site. In Experiment II, we used the video of the Hiro-chan robot

interacting in different manners to collect the participants' opinions. The video consists of a picture of two Hiro-chans sitting on the table, interacting with each other (Figure 4.4). While using two robots, we believe it is important that the participants know which robot is making the voice. During the preparation of the experiment, we tried using voices with different distance based on their position, and different infant tones in the video. However, it was difficult to tell which Hiro-chan was making the voice, even for a younger adult. We believe it would be even more difficult for older adults to tell. For this reason, we applied a highlight to the Hiro-chan that is uttering voices. When the Hiro-chan is making utter emotional voices, a highlighted surrounding appears to suggest which Hiro-chan is making the voice (Figure 4.4 (b) and (c)). The highlight disappears in the pauses between the voices (Figure 4.4 (a)). We understand it is unconventional to highlight the robot since it is not a part of the function of Hiro-chan. But for more accurate experimental results, we made some necessary changes. Figure 5 shows the flow of the video using the example of the time interval of 200ms using both Laugh emotion. Once the first Hiro-chan finishes its voice function, the timer counts 200ms before the second Hiro-chan makes voice. This kind of loop is repeated 10 times. The recorded infant voice is presented in the video. We prepared 24 videos, each representing a different combination of emotional expressions with different voice making pauses. Based on the three emotional expressions explained in Experiment I, we made a combination of emotional expressions using the two Hiro-chans. There are six combinations, which are 1) cry x cry (CC), 2) cry x normal (CN), 3) cry x laugh (CL), 4) normal x normal (NN), 5) laugh x normal (LN), 6) laugh x laugh (LL). Since the order of emotional expressions could potentially affect the result, to avoid differences in the perception of the emotional situation, we kept the same order of emotional

expression in this experiment. The order is the same as above, where laugh x normal means laugh before normal. Since the better voice making pause between humans is 200ms, and the better voice making pause between a human and a robot is between 1s and 2s, we prepared the videos with four turn-taking times, which are 0ms, 200ms, 1000ms, and 2000ms. Since there have not been studies on infant speech turn taking, or that of baby-like robots, we have no other standard to follow in this experiment. Which is why although infants do not follow the speech pattern of adults or social robots, we still use the natural interval of these numbers. We hope to find out the general preference of participants on the speech of baby-like robots first. The six combinations of emotional expressions are made with all four turn-taking times, giving a total of 24 videos in Experiment II.

4.4.3 Measurement

In addition to Questionnaire 1 (see Table 4.1), to understand how older people are affected by the combination of emotional expressions and timing in turn-taking in passive social situation, we added one question:" The two robots are interacting with each other." This question is marked as" Passive Social" in scale. The ratings from 1 to 7 in Experiment I remain the same.

4.4.4 Procedure

The procedure for Experiment II is as follows: The participant remained in the same room when conducting Experiment I. After the participant completed Ex-

periment I, the experimenter opened the questionnaire on the laptop for the participant. Then, the experimenter left the room, leaving the door open, and waited near the room. The participant could ask the experimenter questions whenever he/she had trouble operating the computer or answering the questionnaire. After the participant watched a video and completed the questionnaire, the next set of videos and questionnaire was presented. The order of the videos was counterbalanced among participants. After watching all 24 videos and completing the corresponding questionnaire, the participant called the experimenter. The experimenter confirmed the completion of the questionnaire, and then the experiment was finished.

4.4.5 Results

Table 4.3: Result of Experiment II

Scale	**Cry × Cry (CC)**	**Laugh × Laugh (LL)**	**Normal × Normal (NN)**	**Cry × Laugh (CL)**	**Cry × Normal (CN)**	**Laugh × Normal (LN)**
	M (SD)	**M (SD)**	**M (SD)**	**M (SD)**	**M (SD)**	**M (SD)**
Cuteness	3.97 (0.36)	5.26 (0.38)	5.54 (0.35)	4.33 (0.37)	4.24 (0.33)	5.21 (0.36)
Fun	3.75 (0.36)	5.14 (0.37)	5.35 (0.35)	4.20 (0.34)	4.04 (0.37)	5.13 (0.37)
Interaction	3.99 (0.36)	5.31 (0.37)	5.30 (0.36)	4.27 (0.35)	4.31 (0.35)	5.00 (0.36)
Closeness	3.89 (0.32)	5.07 (0.34)	5.36 (0.31)	4.16 (0.31)	4.16 (0.32)	5.10 (0.33)
Conversation	4.12 (0.37)	5.13 (0.54)	5.54 (0.31)	4.41 (0.36)	4.40 (0.33)	5.21 (0.33)
Passive Social	3.70 (0.37)	5.52 (0.34)	5.13 (0.34)	3.66 (0.37)	3.97 (0.39)	5.30 (0.29)

Table 4.3 shows each emotional expression's meaning and standard deviation in each question. We conducted a three-way mixed-design ANOVA with within-subject factors of the combination of emotional expression and voice making pause and with a between-subject factor of gender on the rating of the questions. There were significant main effects of the combination of emotional expression in all the questions:

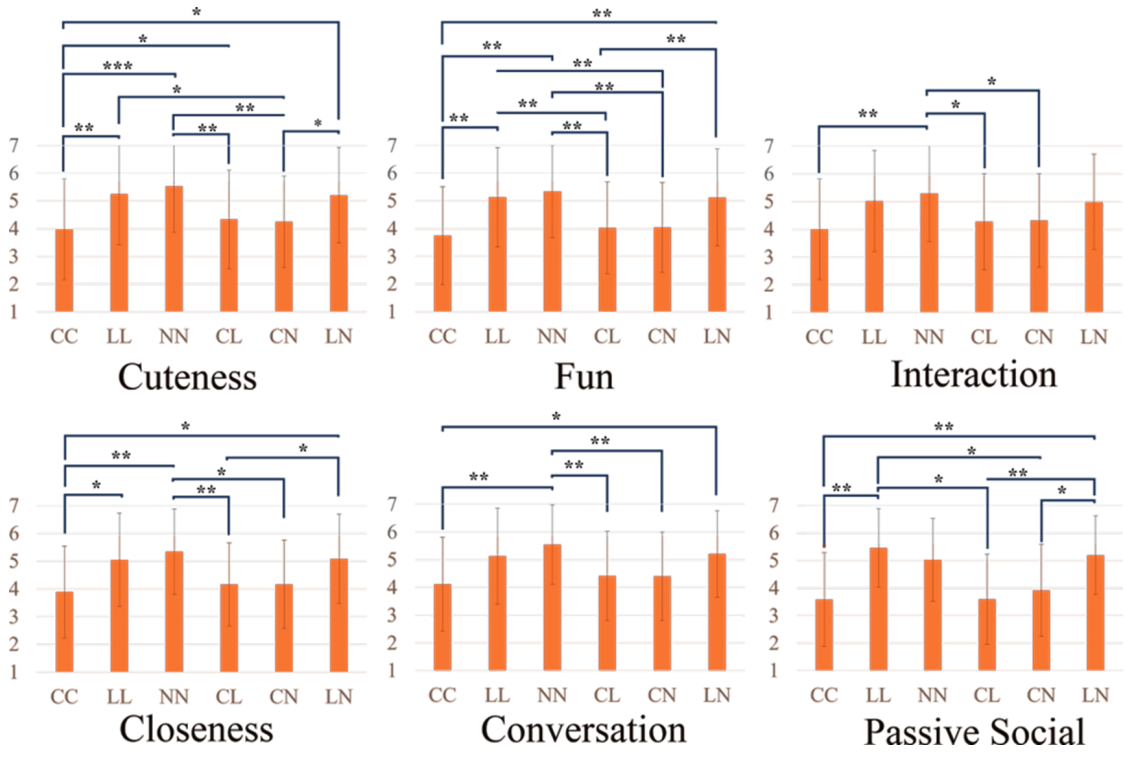
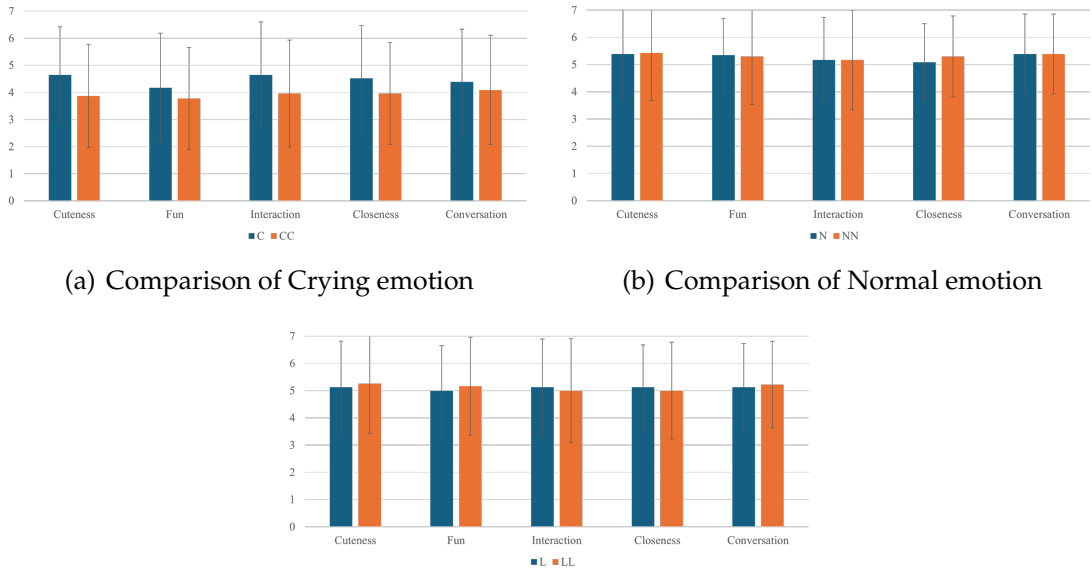


Figure 4.6: The effect of the combination of emotional expressions on each scale ($* : p < .05, ** : p < .01, *** : p < .001$)

Passive Social ($F_{(5,105)} = 11.72, p < .001, \eta_p^2 = 0.28$), Cuteness ($F_{(5,105)} = 15.18, p < .001, \eta_p^2 = 0.42$), Fun ($F = (5, 105) = 19.51, p < .001, \eta_p^2 = 0.48$), Closeness ($F_{(5,105)} = 11.88, p < .001, \eta_p^2 = 0.36$), Conversation ($F_{(5,105)} = 9.63, p < .001, \eta_p^2 = 0.31$), Interaction ($F_{(5,105)} = 8.46, p < .001$), $\eta_p^2 = 0.29$). Figure 4.6 shows all results of the combination of emotional expressions. Post-hoc comparisons with Bonferroni corrections show that the combination NN significantly increases the feeling of closeness, interaction, conversation, cuteness, and fun compared to the combination CN. Combination LL significantly increases the feeling of interaction and fun against combination CL. No significant difference exists among LL, NN, and LN. There were no significant differences in gender.

There was significant main effect of time in turn-taking only for the question 'The robot is cute' with $F_{(3,63)} = 4.83, p = 0.004, \eta_p^2 = 0.19$. Post-hoc comparisons



(a) Comparison of Crying emotion

(b) Comparison of Normal emotion

(c) Comparison of Laughing emotion

Figure 4.7: Comparison between using one and two robots

with Bonferroni corrections show that the feeling that the robot is cute increases significantly in the voice making pause time of 0ms compared to 200ms. (Figure 4.7).

4.4.6 Comparison between experiment I & II

We conducted statistical analysis on the result of experiment I and II to confirm our RQ2: Does using two Hiro-chan robots change the impression of the older people on the baby-like robots? We conducted two sample t-test on the result of using one emotional expression and that of using the same two emotional expressions (e.g. L to LL) for the five scales shared in Experiment I and II (Cuteness, Fun, Interaction, Closeness and Conversation). The result is shown in Figure 8. Though some differences and tendencies can be observed from the result, we found no significant differences.

4.4.7 Discussion

As we predicted, a crying sound inhibited the effect of other emotional expressions, which supports our prediction P2 in all aspects. The combinations which do not include crying sound (LL, NN, and LN) were rated higher than the combinations with crying (CC, CL, and CN). Negative emotions do not seem to have a significant effect on the same emotion since CC does not show significant difference against CN nor CL except for significant difference between CC and CL for cuteness.

We did not find any difference between positive emotion (L) and neutral (N) in their combinations (LL, LN, and NN), which is similar to the result in Experiment I. This implies that the combination of laugh and normal are almost the same in attracting older people using baby-like robots. However, they may work a bit differently. The result passive social scale showed that combination LL and combination LN are significantly higher than all the combinations concerning crying. On the other hand, combination NN was significantly higher than all the combinations concerning crying at the scales of closeness, conversation and interaction. This suggests that combination NN is the better way to attract older people's attention while combination LL should be used to take advantage of the effect brought about by passive social situation. Generally, we find no difference in the time pause of infant voices of baby-like robots. Only cuteness significantly differs in the pause time of 0ms and 200ms. It has been revealed in the study of cuteness, that when two robots talk in synchronicity, people perceive more cuteness than the situation of one robot and the situation of more than two robots [24]. Therefore, we presume this result could be because of the synchronicity of robots, leading to the perception of cuteness. From this result, we can conclude that the timing in turn-taking does not matter in the

infant voices of baby-like robots or that the tolerance of the time pause of baby-like robots is much bigger than that of humans and humanoids. It is curious to find the time in turn-taking where older people find the baby-like robots are not interacting with each other.

Experiment II helped us better understand the emotional expression of baby-like robots. As can be seen from the comparison of Experiment I & II, the result of using one or two robots did not differ significantly (Figure 8). However, this does not mean using two baby-like robots is interaction situation is meaningless. We believe that unlike the social robots with full communication functions, which directly increases the perception of cuteness, baby-like robots do not increase the impression on themselves, but on the interaction as a whole. The differences that could not be observed are obvious in the two-robot situation. From this result, we deduced that using two baby-like robots could enhance the perception of emotional expression, strengthen the impression of the original emotion. In experiment II, for the betterment of distinction of the two robots, we outlined the robot that is making voice functions, while it is not the actual function for Hiro-chan robots. Though we believe the actual robots are better distinguishable even without the outlines, this change may cause better scores in the interaction. In the same experiment, we used a fixed expression order to avoid difference in perception. Though we did not cover this subject in this study, we do believe that different order of emotion could lead to different results, and that this is a subject worthy of further studies.

4.5 General Discussion

From the results of experiment I & II, we have drawn some conclusions. Using one Hiro-chan has very limited effect on changing the impression of older adults. Still, we can see that Crying has negative effects on the will to converse and fun of the robot. This result is consistent with the filed work done with Hiro-chan in nursing homes, where older adults felt annoyed and refused to care for Hiro-chan [14]. In the meantime, though the result of treating baby-like robots remains the same, the reasons behind the result can vary. In the study using Babyloid, it is reported that when it cried, participants felt both positive and negative feelings for the robot, especially at a close distance. In the study concerning real infants, there are also results that suggest complex neuron and hormone changes resulting in an increase in both sensitivity to infant's behavior and stress level [34]. This suggests an infant crying can lead to complex emotions that can be both positive and negative for both adults and the infant. However, unlike in the real infant situation, in the case of a baby-like robot, the positive emotions triggered by crying are overwhelmed by the negative ones, causing a negative result overall. From the result of this experiment, neither the will to interact with the robot nor the will to get close to the robot have increased when robots cry. In the meantime, the Fun and will to converse may be the reason why the impression of the participants decreased in the Hiro-chan experiment. As is generally recognized, the willingness to verbally and nonverbally interact with robots could affect each other, either positively or negatively. However, in the results, we did not observe such tendencies. We believe that it is because during a questionnaire session, emotions do not affect each other as closely as in real situations. Participants tend to analyze each subject

independently when viewing each question. This can lead to some differences with reality, but in this study, we believe viewing each question independently would help with accurate analysis in this study.

4.5.1 Implications

The results of this study can be applied in multiple ways apart from baby-like robots. First, it provided guidance towards the interaction design of baby-like robots. Though we only tested older adults for the sole purpose of lessening loneliness, researchers on baby-like robots will have guidance on how to design their interactions. Secondly, we discovered the method of controlled interaction using baby-like robots. We can change the behavior of robots for different situations and purposes using passive social. For example, making the robots laugh together to attract users, making the robots talk when wanting to keep the user interested, making the robots cry when we want the user to stay away. Though the conclusion is primary, and there are still many more elements to consider, it is a useful method of interaction using baby-like robots. Also, this study enriched the method of human robot interaction. As of now, most of the human robot interaction focuses on verbal or non-verbal communications. Using baby-like robots explored the territory of communication using emotional expressions. While some may consider not being able to use words is a limitation to baby-like robots, the unique trait of infants makes it possible to communicate using purely sounds. And unlike animal-like robots, baby-like robots may be considered intelligent and thus more efficient interactions may be possible and deeper connections may be built between the robot and the user. Through this study and further studies, communication using emotional

level can be more systematically explored. Last but not least, through this study, we found a method to study infant-related research topics safely and ethically. There have been studies on the effect of infants on adults, such as the effect of crying or caregiving behavior on adults. But there are still many unstudied subjects in infants and as an uncontrolled study subject, forcing infants to do something could be considered unethical. With the methodology of this study, it is possible to study infants without ethical problems.

4.5.2 Limitation

This study has several limitations. First, due to comprehensive reasons, we used video instead of real robots in this experiment. This could cause some difference in the result. Now that we have confirmed the impression of older people on all six scales, and that the crying stimuli will not disturb the older people too seriously, we can continue on experimenting with real Hiro-chan robot in nursing home in the future. Second, sample size is relatively small. We need to increase the number of participants to confirm current results. Also, during the experiment, though we took many possible precautions to make sure the older adults can complete the questionnaire using a tablet, the task of watching 24 videos and completing the questionnaire after each video was still challenging. As mentioned, one of the participants failed to complete the task. The heaviness and boredom of the task, as well as the similarity of the videos may affect the result in a bad way. However, to minimize this effect, during the experiment, the physical and mental conditions of the participants were monitored through a camera by the experimenter. Though this, we noticed one participant could not complete the task. Since baby crying and laughing are both sharp voices,

it may be difficult to tell the difference, especially for older adults. In the interview after the experiment, few participants reported this. Though only a very small percentage of participants reported this effect, and we do not think this would largely affect the result, we consider this as a limitation and will try to improve in further experiments. In the video used in the experiment, we used an image of Hiro-chan positioned on the edge of the table, which is a rather dangerous position for an infant to be at. This could cause the scales of Closeness and Interaction to be higher. However, we do not believe that this could affect the result of this study, which is to study the difference between emotions. It is worth further study whether the position would affect the impression and how changing positioning could be beneficial to the human-robot interaction using baby-like robots. It is a very important subject to study in the future. In the experiments, we prioritized the quality and the quantity of the audio files but did not control the length of each voice segment the same. For this reason, there are some time differences for each kind of emotion with all the lengths of audios added in total (Laugh= 17.0 sec, Cry= 12.7 sec, Normal= 10.4 sec). This may affect the result in some ways. There has been research on the influence of length on people's focus in both audio and video, but the research mainly focuses on either linguistic audios, educational videos, or background music. In the meantime, the difference usually occurs by minutes instead of seconds. That is why we consider the length to be less influential in the experiment. However, we cannot deny the possibility of an influence on this subject. Expressing emotions for too long may bore people. It would be interesting to investigate how differences in the duration of a robot's emotional expressions affect people's impressions. Finally, further research is needed to determine if our results can be generalized to younger people and older adults with dementia, since we focus

on healthy older people.

4.6 Conclusion

In this study, we researched the better design of voice function in baby-like robots to relieve the loneliness problems of older people in nursing homes. We found what emotional expression attracts older people better and what emotion makes older people not want to use robots. Though the primary purpose of this study is to find out what kind of emotional expression attracts older people better, the result of this study can be extended. By applying preferred emotional expressions to attract older people and applying disliked emotional expressions when the baby-like robot is damaged or used for too long, precise control of human-robot interaction could be achieved. By the result of this experiment, we fill the gap of the voice and expression design of baby-like robots for the caring of loneliness of older people. We believe this study can provide instruction on how to alleviate loneliness and improve the well-being of older people through baby-like robots. This study also provided instruction on the future work for Hiro-chan. In the future, we plan on extending this study by using actual Hiro-chan robot in nursing home.

CHAPTER 5

PURPOSE DESIGN OF HIRO-CHAN

In Chapter 2, 3 and 4, we conducted the morphological design and functional design of Hiro-chan and made conclusions on the psychology behind each design. As a further step, we would like to explore the new interactive method using Hiro-chan. In the study conducted in caring facilities using Hiro-chan, it was observed that communication among older adults as well as the communication between older adults and the care staff increased with the presence of Hiro-chan. Studies have shown that communication with other human beings can improve the health conditions of both physical and mental health. This is where we were provided with the idea that Hiro-chan can facilitate communication between older adults. In the functional design, we found different emotions of Hiro-chans can have different effects on the reactions of older adults. In the purpose design, we may be able to utilize the conclusions and use those emotions to induce communication. By applying Hiro-chan as a media to connect older adults together, we can benefit the health of older adults better than merely having Hiro-chan interact with older adults alone. This study can discover the new method of human-robot interaction using baby-like robots where the robot serves as a media, achieving a human-centered human-robot interaction. In this chapter, we conducted the experiment using two Hiro-chans with two older adults each having one Hiro-chan near them to see how Hiro-chans affect the relationship and communication between the older adults. We found that the relationship of the older adults largely depends on the psychology of the older adults, while the communication depends on both the psychology and the emotion of Hiro-chan.

5.1 Background

5.1.1 Interactive social

Interactive social in robot-mediated settings conceptualizes the robot not as a conversational partner but as a social medium that structures interaction between two people. By emitting timed, legible, low-intrusive cues, the robot scaffolds joint attention, lowers entry costs for speaking, shapes turn-taking, and helps participants build common ground. Through arousal/salience modulation, such cues manage attention: low-arousal, predictable signals function as a conversational background that increases willingness to talk, whereas high-salience signals may redirect focus to the robot unless jointly appraised, limiting relational gains. The impact is person-contingent and culturally shaped by display norms. Practically, interactive social offers a lightweight social catalyst for warming up dyadic conversation and—when paired with brief shared meaning-making—can be extended to strengthen interpersonal closeness, especially in short sessions with older adults.

5.1.2 Balance theory

Balance theory, originally proposed by Heider [13], explains how individuals strive for cognitive consistency in their social relations. The theory suggests that people are motivated to maintain balanced triadic relationships, in which attitudes toward other persons and objects are consistent and harmonious. When imbalance occurs, such as when two individuals hold contradictory evaluations of a third entity, psychological tension arises, motivating them to adjust their

attitudes or behaviors to restore balance. This framework has been widely applied in social psychology to understand friendship formation, attitude change, and interpersonal attraction. In the context of our study, balance theory provides a useful perspective on how an interactive social robot can serve as a mediating third party between two human participants. By engaging both individuals with the robot in a positive and consistent manner, the robot may facilitate balanced triadic relationships, thereby enhancing the perceived affinity between the participants and increasing their willingness to communicate with each other.

5.1.3 Social capabilities and social anxiety & social self-esteem

Social anxiety, social self-esteem, and social competence are closely intertwined constructs in the field of interpersonal functioning [14][15]. Social anxiety is characterized by heightened fear and avoidance of social interactions due to concerns about negative evaluation, which often undermines individuals' willingness to engage with others. In contrast, social self-esteem reflects individuals' perceived worth and confidence in social contexts, serving as a psychological resource that can buffer against anxiety in interactions. Lower levels of social self-esteem are frequently associated with higher social anxiety, creating a cycle in which diminished self-regard fuels avoidance behaviors, further restricting opportunities for positive social experiences. Social competence, or the ability to communicate effectively and maintain relationships, is both a consequence of and a protective factor against this cycle. Individuals with well-developed social skills are better able to navigate challenging situations, which can reduce social anxiety and reinforce self-esteem through successful interactions. Con-

versely, deficits in social competence may exacerbate the negative effects of social anxiety and hinder the development of social self-esteem. Taken together, these dynamics highlight the importance of interventions that target all three dimensions simultaneously, aiming to reduce social anxiety, strengthen social self-esteem, and cultivate practical communication skills.

5.2 Current Study

5.2.1 Hypothesis

In this study, we aim to find out how using two Hiro-chans influences the relationship and the communication between two older adults. On the base of the knowledge already acquired, we make the following hypothesis.

H1: Cry emotions of two Hiro-chans would increase the relationship between two older adults and the impression of the older adults on each other.

Based on the previous studies where Hiro-chan crying resulted in decrease of impression and positive interactions of older adults [7], we believe that Hiro-chan would have the same effect in the interactive social situation.

H2: Normal emotions of two Hiro-chans would increase motivation to communicate between two older adults.

Based on the balance theory, we know that if the two participants dislike the same thing, their relationship would increase. Since we made the conclusion in Chapter 3 that Cry emotion is the most disliked emotion, we assume that having a shared target for dislike in interactive social situation would increase the relations between two older adults.

Since in Chapter 3, we made the conclusion that Normal emotion would keep older adults longer in the conversation, we believe that in the interactive social scenario, this emotion would also facilitate the communication between the two older adults.

Study shows personal characteristics, such as social self anxiety level and social self-esteem level largely affect quality of social interactions[14][15]. Since the social interaction between people largely depends on the characteristics of the people involved, we must include psychological elements in this study. This is why we also made the following research questions.

RQ1: How is social anxiety related to the effects of using two Hiro-chans as media?

RQ2: How is social self-esteem related to the effects of using two Hiro-chans as media?

5.2.2 Participants

In this experiment, we recruited 24 participants. Within them, 12 were male and 12 were female. All of them were older adults over the age of 60, under the age of 00 ($M = 72.04, SD = 5.83$). The avergae age of male participants is 75.08 (SD= 5.16), with average age of female participants being 69.00 (SD= 4.78). All of the participants were healthy adults with no symptoms of dementia. All the participants were native Japanese speakers. The participants were recruited from a human resource center. We did not set any screening standards except for the age, health condition and gender group mentioned above.

5.2.3 Apprautus

In this experiment, we aim to understand how the different emotions of Hiro-chan would influence the interaction and relationship between two older adults during an interaction. We used two Hiro-chan robots with different clothing to distinguish them. The Hiro-chans used in this experiment had controlled voice functions. The voice functions contain three different emotions, which are Laugh, Normal and Cry. Laugh and Cry are the recorded laughing and crying sound of babies. Normal is the recorded voice of the bubbling sounds of babies. The Hiro-chans were designed to make uttered voices at destinated time. The timeline of the voice functions is shown in Fig 5.1. The timeline is the same for all three emotions. The rest of the functions of the Hiro-chans used in the experiment were the same.

5.2.4 Measurement

In the experiment, we took mainly three types of data. We took questionnaires, the behavioral analysis of the participants and the interview of the participants. The questionnaire includes three parts. The social anxiety test, the social confidence test and the questionnaire in experiment. We use the social anxiety questionnaire for adults[16]. It contains 30 questions and is proven valid for measuring the social anxiety level of adults. Since the original questionnaire was in English, we translated the questions into Japanese. For the social self-esteem questionnaire, we used the Japanese shortened version of Texas Social Behavior Inventory[17]. It measures social self-esteem using 16 questions. For the interactive questionnaire, we asked three questions. 1) IOS (Inclusion of the

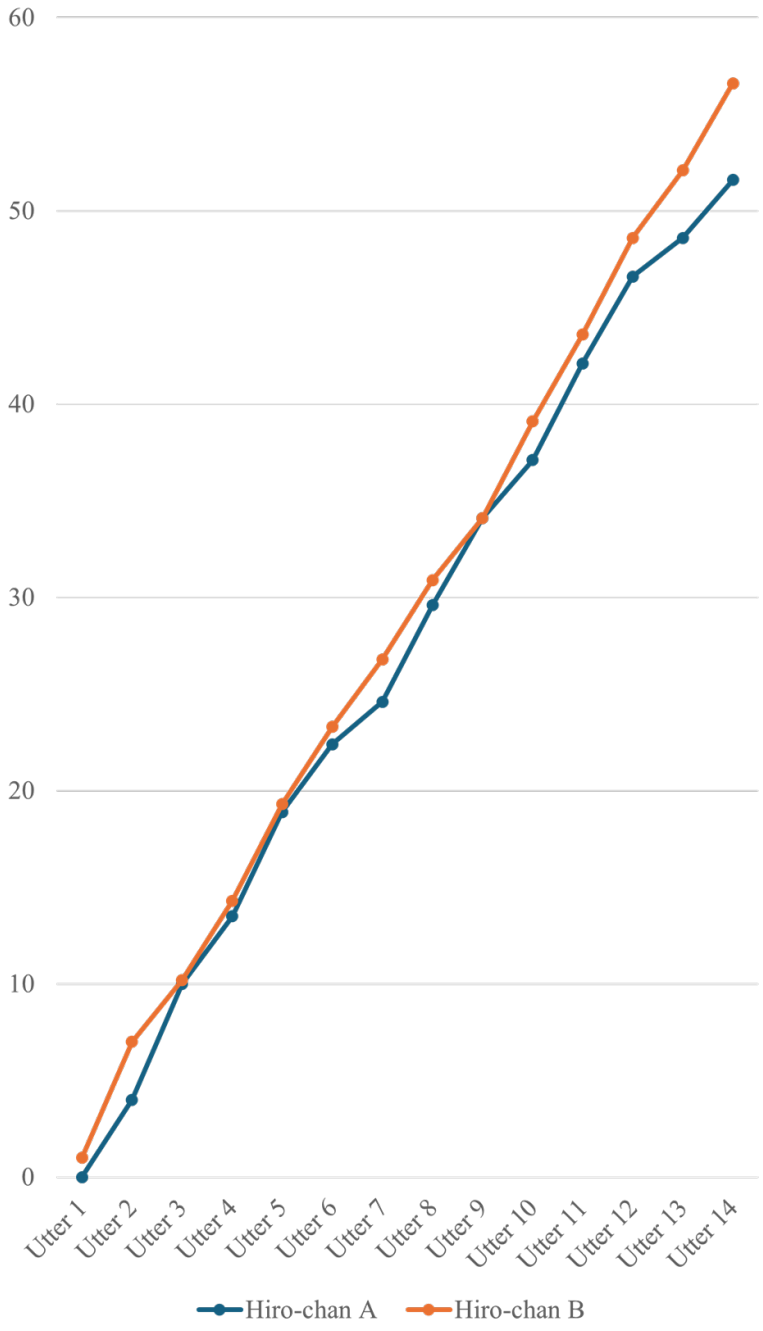


Figure 5.1: Timeline of the voice utterance of Hiro-chans

Other in the Self) test[18], 2) Rate the impression on the other participant, 3) Rate the willingness to communicate with the other participant. Each question had scores from 1 to 7, where 1 is the lowest and 7 is the highest. To under-

stand the change made by the interaction using Hiro-chans, we conducted the questionnaire before and after each interaction.

Procedure

After the participants arrived, we first took the consent from all of the participants. After the consents were taken, we took the social anxiety and the social self-esteem test of each participant. The participants were marked as A, B, C and D. Then the participants were separated into groups of two. The first two participants were led into the experiment room by the experimenter, inside were a table with two Hiro-chans on each side. The participants then took seats with one Hiro-chan in front of each of them. The experimenter then explained the contents of the experiment and took the interactive questionnaire Q1 before the experiment. The experimenter explained if the participants had any questions. After the questionnaires were taken, the experimenter left the room, letting the participants interact with Hiro-chans and each other for the duration of the voice functions of the Hiro-chans, approximately two minutes. During the interaction, the experimenter was observing the experiment through a camera, making sure nothing was wrong or dangerous for the participants. After the interaction, the experimenter entered the room and took the interactive questionnaire. The same experimental process was repeated 6 times, until all four participants were together in the interaction of the group of two (AB, AC, AD, BC, BD, CD). The emotions of Hiro-chan were designed so that all the participants would experience each emotion once. Furthermore, since we conducted the experiment with 12 male and 12 female participants, we made sure the order of emotions of Hiro-chans were different each time.

5.2.5 Analysis

Social anxiety test

For the results of the psychological test before the experiment, we processed the data so that we can estimate the social anxiety and confidence level of the participants[16]. In the social anxiety test, we calculated the average score of each participant. We classified the level of social anxiety into three levels. The highest-level scores ranged from 3.4 to 5. The middle level ranged from 1.7 to 3.3. The lowest level ranged from 1 to 1.6 points. We found the average social anxiety score of all participants to be 2.9, which is the middle level.

Social self-esteem test

In the social self-esteem test, we followed the analysis process of the Japanese shortened version of Texas Social Behavior Inventory. We calculated the social-withdraw score and the social confidence scores of participants, men and women respectively using different questions, then classified them into high social self-esteem group and low social self-esteem group[17]. In the men participants, 3 participants were classified in high social self-esteem group and 9 were classified in low social self-esteem group. In women participants, 3 of them were classified in low social self-esteem group, and 9 of them were classified in high self-esteem confidence group.

Questionnaire results

For the questionnaire results, we subtracted the scores after the interaction to the scores before the interaction for all three questions respectively. Then we conducted a four-factor ANOVA on the results with the within-subject factor being the emotion of the Hiro-chans (laugh, normal, cry) and three between-subject factors of participant gender (men, women), social anxiety level of the participants (high, middle, low) and the social self-esteem level of the participants (high, low). We conducted the same analyzation for all three questions (Ios test, impression on the other participant, willingness to talk to the other participant). For all the analyzation, we first applied Mauchly's test to confirm an assumption of sphericity. If it was violated, the degrees of freedom were corrected using the Greenhouse-Geisser estimates of sphericity. We conducted post-hoc pairwise comparisons with Bonferroni corrections.

5.3 Results

IOS

Figure 5.2 shows the results of Ios test. We found significant differences in the main effect of Ios test scores in social self-esteem level ($F_{(1,14)} = 8.93, p = .010$, $\eta_p^2 = .39$) and participant gender ($F_{(1,14)} = 6.55, p = .023$, $\eta_p^2 = .32$). We also found significant difference in interaction between social self-esteem level and participant gender ($F_{(1,14)} = 28.12, p = .007$, $\eta_p^2 = .42$). Post-hoc comparison with Bonferroni corrections showed that the Ios scores of women low social self-esteem group ($M = 4.17, SD = 1.20$) increased significantly more than that of

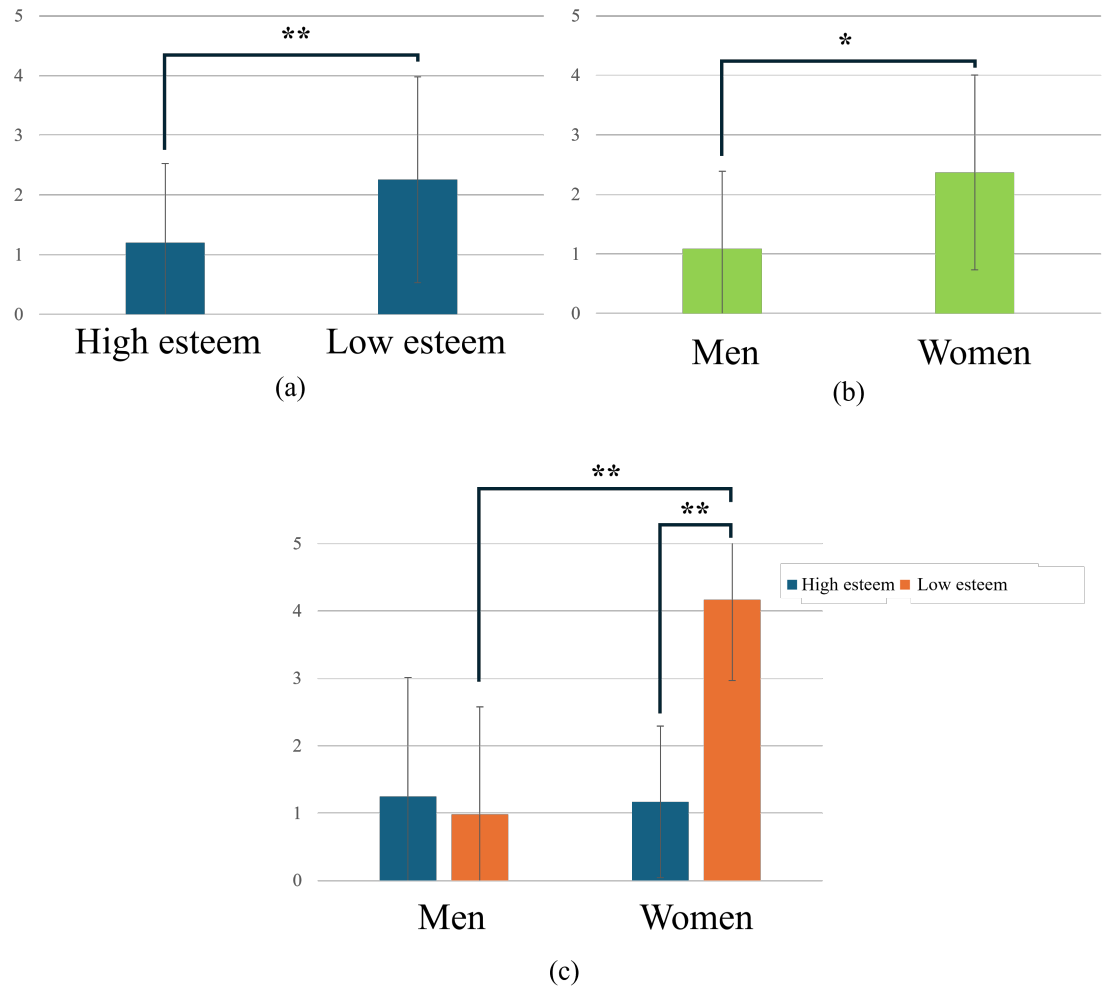


Figure 5.2: (a) Increase in scores of Ios test against social self-esteem level; (b) Increase in scores in Ios test against participant gender; (c) Increase in scores of Ios test against interaction between social self-esteem level and participant gender ($* : p < .05, ** : p < .01$)

men ($M = 0.98, SD = 1.59, p = .001$). In female participants, low social self-esteem ones had significant higher score increases than high social self-esteem ones ($M = 1.17, SD = 1.12, p = .001$).

Impression

We did not find any differences in the increase of scores in the impression on the other participant.

Willingness to talk to the other participant

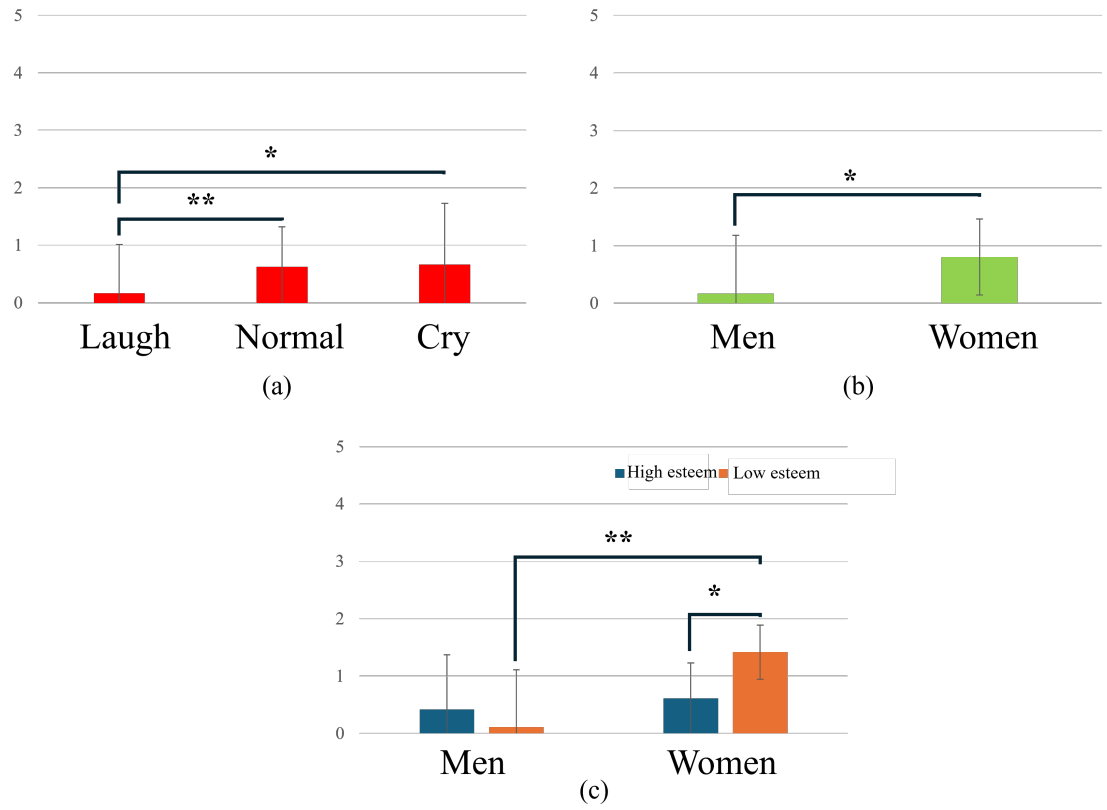


Figure 5.3: (a) Increase in scores of willingness to talk to the other participant among robot emotions; (b) Increase in score of willingness to talk to the other participant against participant gender; (c) Increase of score of willingness to talk to the other participant against the interaction between social self-esteem level and participant gender($* : p < .05, ** : p < .01$)

Figure 5.3 shows the results of the question of the will to talk to the other participant. We found significant differences in main effects of increase in questionnaire scores in robot emotion ($F_{(2,28)} = 4.90, p = .015, \eta_p^2 = .26$) and partici-

pant gender ($F_{(1,14)} = 8.34, p = .012, \eta_p^2 = .37$). We also found significant difference in the interaction between social self-esteem level and participant gender ($F_{(1,14)} = 5.20, p = .039, \eta_p^2 = .27$). Post-hoc comparison with Bonferroni corrections showed that in the low social self-esteem group, female participants ($M = 1.42, SD = 0.47$) showed significant more increase in scores than male participants ($M = 0.11, SD = 1.00, p = .005$). Also, for female participants, low social self-esteem ones showed significant increase in scores than high social self-esteem ones ($M = 0.61, SD = 0.62, p = .049$).

5.4 Discussion

5.4.1 Implications

We investigated the effect of different emotions of two Hiro-chans in the interactive social situation on the relationship and communication between two older adults. The result of the questionnaire results suggested that the increase of relationships between participants is not biased on the emotions of the Hiro-chans. The older adults in the interaction are not dependent on the emotions of Hiro-chans, but on the social self-esteem level and the gender of the participants. From these results, we can say that our H1 is not supported. At the same time, we can see that the willingness to communicate with the other participants varies by the emotions of Hiro-chan. Having normal and cry emotions making significantly more increase in scores of willingness of communication than laugh emotion, we can say that our hypothesis H2 is supported. Also, after investigating the social anxiety and the social self-esteem of the participants

and combining the results with the questionnaire scores, we had some interesting findings. First of all, we found that in the IOS score, which illustrates the relationship between the participants, the increase in scores is more obvious in participants with lower social self-esteem, which are the people who are more socially withdrawn and less socially confident. There have also been some interaction effects between gender and social self-esteem level, suggesting that interactive social situation using Hiro-chan can have different effects on people with different psychology and different genders. The same can be said for willingness to communicate. These results can bring some interesting conclusions for our research questions.

5.4.2 Differences in emotions of Hiro-chans

Unlike what was expected, we did not find many differences in the increase of scores except in the willingness to communicate. We found that in the in the willingness to communicate, the increase of scores after the interaction is greater in normal and cry than in laugh. We believe that the reasons for the significant increase in normal and cry are different. For the normal emotion, we believe that it is because of the trait of normal emotions that keeps older adults in the interactions. As expected, we found in the functional design that normal emotion can increase the will of users to interact with Hiro-chan for a longer period of time and the will of users to talk with Hiro-chan. For the same reason, in the interactive social situation, participants felt the desire to interact and talk and projected the desire to the other participant as well. For the crying emotion, we believe that this is due to the effect of a combined effect of balance theory and shared adversity. Studies have shown that shared adversity can improve team

coherence and creativity. Balance theory points out that people build deeper bonds if they face something they both have negative impressions on. In this situation, Hiro-chans crying is a dislikeable situation for both of the participants and they have to endure shared adversity. This caused the participants to build more communications either to get through the adversity together or to ask advice from the other participant. Either way, the will to communicate is increased in this situation. The reasons for the increase in normal and cry emotions are different and we believe that is the reason that there was no significant difference between normal and cry emotions.

5.4.3 Differences in gender

In the results, we found that gender is a significant main effect for both IOS and willingness to communicate. This suggests a difference in the interactive social situation between genders using Hiro-chan. Hiro-chan may work better among female users than male users.

5.4.4 Differences in social self-esteem level

We found difference in increase of scores of IOS with the social self-esteem level. Participants with lower levels of social self-esteem felt closer to the other participant after the interactive social situation with Hiro-chans. In the meantime, the interaction of gender and social self-esteem level is found significant in both IOS and willingness to communicate, giving the result that female with lower social self-esteem endured more increase in relationship and communication.

We believe that this is a meaningful result that people who do not actively socialize can experience social bonds as well as the will to create social bonds with the presence of Hiro-chan. The difference in the increase of scores is due to the people with lower social self-esteem using Hiro-chan as a media to connect with the person besides them, while the people with higher social self-esteem do not need the media to communicate. We believe that this result suggests those participants crave social interactions, while they are either not good at socializing or not confident enough to start a conversation with the subject besides them, and that Hiro-chan in this experiment served as a media connecting the participants. We believe this is proof that Hiro-chan as a media connecting older adults is a effective method.

5.4.5 Limitations

This study has some limitations. First of all, since the main objective of this study is to understand the effect of different emotions and psychologies on interactive social situations, we did not compare the social interaction with and without Hiro-chan. Thus, we cannot make the comparison between mere socializing and socializing with Hiro-chan as a media. Secondly, in this study, the main evaluation method is the questionnaire. It would be better and more accurate with the behavioral and speech analysis of the recording of the experiments. With the speech analysis, we could better understand the results of the questionnaire. Thirdly, we believe it is possible to dig deeper into the combination of participants with different psychological traits in the interactive social situations with Hiro-chan. For example, it would be interesting to study situations with two extrovert participants, one extrovert and one introvert participant and

two introvert participants using Hiro-chan as a media.

5.4.6 Conclusion

This study examined how deploying two Hiro-chans as a social medium shapes dyadic interaction among older adults. We set out to test whether robot emotions would differentially influence relationship and communication, and whether person factors would moderate these effects. The findings are clear on two points. First, H1 was not supported: robot emotion did not improve IOS or partner impression; closeness was instead explained by social self-esteem and gender, with the largest IOS gains observed among women low in social self-esteem. Second, H2 was supported for communication: emotion affected willingness to talk, with Normal and Cry conditions producing larger increases than Laugh. Notably, Normal and Cry yielded similar gains in willingness, likely via different mechanisms—Normal acting as a low-arousal scaffold that lowers entry costs to conversation, and Cry creating a jointly recognized, mildly aversive context that can prompt coordinating talk—without translating into deeper relationship change within a brief session. Across outcomes, social self-esteem emerged as a robust moderator: participants lower in social self-esteem—especially women—showed the strongest increases in IOS and willingness to talk, consistent with the idea that predictable, legible cues provide valuable conversational scaffolding for those least confident about initiating social contact. By contrast, social anxiety did not show reliable moderation in this short, externally structured setting. Practically, these results position interactive-social use of robots as a lightweight social catalyst: if the design goal is to "get them talking," favor low-arousal, predictable cues; if us-

ing Cry, pair it with brief, shared meaning-making to convert aversive salience into aligned interpersonal engagement; avoid Laugh when the immediate aim is to raise partner-directed talk. Personalization should prioritize low social self-esteem users, for whom the benefits are greatest. Taken together, the study demonstrates that robots used as media—rather than partners—can increase communication intent in older-adult dyads while relationship change remains person-contingent and unlikely to shift from brief exposure alone. Future work should add behavioral speech/gaze measures, include a no-robot baseline, extend to multi-session protocols, and apply dyad-aware models to clarify how robot emotion and individual differences jointly produce durable social benefits.

CHAPTER 6

GENERAL DISCUSSION AND CONCLUSIONS

6.1 Discussion

6.1.1 Practical significance and design guidance

In the morphological design, baby-like robots should keep an identifiable infant body outline with torso and limbs and use infant voice. This combination increases caregiving actions, reduces non-caregiving behavior, and provides a stage for conversation between older adults. In the functional design, baby-like robots should use bubbling sound of infants when the goal is to get people to start talking, use Cry only in short doses and with shared meaning making if possible, so that it becomes a moment of joint response rather than a burden. Using baby-like robots provides the possibility to shift the purpose from human-robot interaction to using the robot to support human-human interaction. The robot works as a low-effort social catalyst that can ease caregiver workload and lift the social atmosphere among residents.

6.1.2 Methodological and theoretical contributions

The study connects the multimodal infant-schema mechanism to elder care robots and validates the boundary of minimal design using both subjective ratings and behavioral indicators. It links short-term emotional acceptability to effects on human-human interaction and positions the robot as a social medium.

This offers a fresh approach for group-oriented social interventions and provides clear evaluation paths.

6.1.3 Application and future work

This thesis offers a general way to use a baby-type robot as a social medium that helps older adults talk with each other in everyday settings. With a clear and gentle body shape and a soft infant voice, the robot provides simple cues that people can notice together without feeling pressure. In short moments when two people meet and do not know how to start, these cues can create a shared focus and make the first words easier. During routine times, the same cues can keep a calm rhythm so that talk feels smoother and lasts longer. In rooms with several pairs, two or more robots can lift the overall mood and raise the density of small conversations. Around daily transitions such as before meals or before group activities, the robot can act as a soft bridge that helps people move from silence to light talk. Because the cues are easy to understand and do not depend on language, this approach is inclusive for a wide range of users and settings. It aims to support human connection rather than replace it, brings low burden for staff, respects privacy by avoiding personal data, and can be scaled across rooms and routines. Future work should show how short gains in starting talk can grow into stable relationships and then into benefits for well-being. The focus can move from a single pair to the social network of a whole room or a whole facility, so that we learn how local changes spread through a group. Design research can build a simple shared language across sound, small motions, gaze and gentle light, with clear rules for timing and strength that work in many spaces. The system can learn to adjust volume and cue density to the noise

and activity level in the room so that it stays helpful without drawing attention away from people. Evaluation can bring together brief self-reports, basic measures of talk and turn taking, and simple room level indicators, so that results are comparable across sites and over time while keeping privacy safe. Ethical guidance should be made clear and easy to follow, with transparent notice and data minimization. Robots can be linked with other non-drug activities such as reminiscence, singing, exercise and crafts to look for combined gains. Studies across different cultures and care models can test how general the approach is and what parts need local tuning. Finally, theory can advance by building models that connect cue settings to attention, to conversation, and then to social outcomes, while real world work can test cost, sustainability and staffing benefits for long term use. In this thesis, Hiro-chan was used to aid older adults with cognitive degressions in their loneliness issues. The method is to use emotional interaction with Hiro-chan to help older adults that have difficulties in cognitive interactions. In the future, I believe such study can be expanded further in the people with problems in cognitive interactions. It is generally believed that the relationships between two people are built upon cognitive interactions[19]. Such process is introduced as social penetration theory[20]. However, for those people who have difficulties performing cognitive interactions, socialization can be difficult for them and thus loneliness, social withdraw happens. In the study using Hiro-chan targeting older adults with cognitive degression, emotional interaction is effective in the one-on-one interaction with them, as well as in two-on-two interaction using two Hiro-chans and two older adults. Participants showed communication based on emotions with the other participant according to the behavioral analysis. I believe Hiro-chan has the potential to induce emotional communication between people who do not already have a

close relationship, or even those who first met. By using Hiro-chan, we can alter the traditional way of building a relationship from cognitive interaction to emotional interaction and discover how to start from emotional interaction to cognitive interaction. Such process is used in the elderly-to-elderly communication in purpose design. In the future, once we have understood how to use Hiro-chan to build relationship of emotional interaction to cognitive interaction with older adults, we can expand the application further to see if the same process works with relationships between different age groups, such as between the older adults and the care staff. Such relationships can benefit both the older adults and the care staff in the caring facilities. Also, such study can be expanded to include all kinds of people who have trouble socializing for the reason of incapability of cognitive interactions, such as children with less developed cognitive functions, or people with autism spectrum disorder (ASD). With this study, people having trouble conducting cognitive interactions can be benefit and the application of Hiro-chan can be expanded to more than elderly care. This study was conducted in the context of Japanese elder care. It responds to two practical needs in facilities: limited staff resources and the mental health of residents, including loneliness and behavioral and psychological symptoms of dementia. We examined how a baby-type social assistive robot, Hiro-chan, can better support older adults. The work followed a three-part program across morphology, function, and interactive social design. The goal was to identify design choices that are easy to use, ethically sound, and effective for day-to-day care.

6.2 Conclusion

We aimed to understand the psychological mechanisms that make Hiro-chan effective and to turn those mechanisms into concrete design choices. We organized the work into three steps. First, we studied morphology to define the minimal but sufficient baby-like features. Second, we studied function to clarify how different emotions, Laugh, Normal, and Cry, shape user experience and long-term acceptance. Third, we studied interactive social settings to test whether using two Hiro-chans can promote communication and relationship between two older adults. When we compared different forms, the baby-form with identifiable torso and limbs was rated the most baby-like and was clearly above the other shapes. The presence of voice and participant gender also affected baby-likeness ratings. In behavior, the baby-form elicited the most caregiving actions, and adding an infant voice further increased these actions. Non-caregiving behavior was lowest for the baby-form, and the presence of voice also reduced such behavior. For infant directed speech and non-infant directed speech, voice was the main driver, while form and gender were not primary factors. Together these results show that minimal does not mean removing key information. To reliably trigger a caregiving script, a clear, identifiable infant body outline combined with an infant voice is an efficient, low-load design. The voice can compensate when the visual form is abstract, which strengthens the infant attribute and supports empathic imagination. Through the results found, an inferred model between infant features in baby-like robots, baby schema and caregiving behavior is built, and the general guideline for minimal design of baby-like robots is set to fill the blank in this research area. However, further study is needed to determine the statistically accurate model, as well as the

exact level of minimal design instead of the generally one we concluded in this thesis. Prior evidence suggests that adults respond to infant laughter and crying through reward, alerting, and caregiving systems, while long exposure to strong negative cues can cause fatigue and avoidance. Field feedback indicated that keeping crying as a frequent cue lowers positive impressions and willingness to use the robot over time. This supports a functional choice to control or limit Cry and to prioritize emotional expressions that remain acceptable during long-term use. In dyadic sessions with older adults, we tested relationship measures and willingness to talk. Hypothesis 1 was not supported. Robot emotion did not raise IOS or partner impression. Relationship gains were explained by person factors, especially social self-esteem and gender, with the largest gains among women with low social self-esteem. Hypothesis 2 was supported for communication. Both Normal and Cry increased willingness to talk more than Laugh, and their increases were similar in size, although this did not convert into relationship change within a short session. The two emotions worked through different paths. Normal provided a low-arousal, predictable scaffold that lowered the effort to start talking. Cry created a mild shared difficulty that prompted coordinating talk. Results also showed an interaction of gender and social self-esteem. Women with low social self-esteem improved the most on both IOS and willingness to talk. This indicates that personal characteristics determine who needs and benefits most from gentle social support.

ACKNOWLEDGMENT

As I finish my three-year doctorate program, I would like to express my gratitude towards the people who supported me throughout the process.

Firstly, I would like to pay my gratitude towards my parents. They have been nothing but supportive for me both financially and mentally. Without their belief in me, I could not have accomplished this.

I want to pay my thanks to my professor, Hiroshi Ishiguro and my supervisor Hidenobu Sumioka. They guided me into the field of my study, providing suggestions and ideas during the study and giving me instructions every step of the way. Their instructions gave me inspirations towards experiments and study directions. I would also like to thank Sayuri Yamauchi, Yasue Miyazaki, Yukiko Takei and all the members in ATR assisting in the experiments and paperwork towards conferences. Their assistance was necessary and helpful for this study. I also appreciate the help from Doctor Uchida and the secretary of Ishiguro Laboratory, Oshima-san, for their help in the process of this thesis.

I would like to thank my girlfriend Siyu Yang for accompanying me through the doctorate life, giving comfort and confidence all the way. I feel lucky to have met her and to have the courage to express my feelings. Her support was essential in my study and I hope my work is worth the wait. I hope I can give her the same support in the future in her hard times and hope that we can be of each other's help for the rest of our lives.

Thanks to all my friends for their emotional support and pressure relief through games. Special thanks to Haoyu Liu for his support in my knowledge of ANOVA.

I also want to thank Marie Morimoto for being cute, high spirited and encouraging at live concerts and meet & greet times. It was always fun to see and talk to her at events. Her encouragement gave me energy towards completing the dorctoral research.

Finally, I want to thank Rena Moriya for being gorgeous and always showing the best smile. The thought of seeing her smile in live and meet & greet events motivated me in the doctoral study.

BIBLIOGRAPHY

[1] M. Sato, Y. Yasuhara, K. Osaka, H. Ito, M. J. S. Dino, I. L. Ong, Y. Zhao, and T. Tanioka, "Rehabilitation care with Pepper humanoid robot: A qualitative case study of older patients with schizophrenia and/or dementia in Japan," vol. 30, pp. 32–36.

[2] T. Shibata and K. Wada, "Robot Therapy: A New Approach for Mental Healthcare of the Elderly – A Mini-Review," vol. 57, no. 4, pp. 378–386.

[3] Z. Shen, N. Urano, C.-P. Chen, S. Feng, S. Mitchell, M. Katagiri, Y. Du, F. P. Trevejo, T. P. Tomo, A. Schmitz, R. Browne, T. Ogawa, Y. Taki, and G. Trovato, "Participatory Design and Early Deployment of DarumaTO-3 Social Robot," in *Social Robotics* (F. Cavallo, J.-J. Cabibihan, L. Fiorini, A. Sorrentino, H. He, X. Liu, Y. Matsumoto, and S. S. Ge, eds.), pp. 550–559, Springer Nature Switzerland.

[4] A. Nakae, H. M. Bu-Omer, W.-C. Chang, C. Kishimoto, and H. Sumioka, "Exploring the Psychological and Physiological Effects of Operating a Telenoid: The Preliminary Assessment of a Minimal Humanoid Robot for Mediated Communication," vol. 24, no. 23, p. 7541.

[5] H. Sumioka, N. Yamato, M. Shiomi, and H. Ishiguro, "A Minimal Design of a Human Infant Presence: A Case Study Toward Interactive Doll Therapy for Older Adults With Dementia," vol. 8.

[6] W.-Y. Chou, P. J. Clarkson, and C. Waszynski, "It Still Matters: Preference in Visual Appearance of Stimuli Among People in the Late Stages of Dementia," in *Human Aspects of IT for the Aged Population. Acceptance, Communication and Participation* (J. Zhou and G. Salvendy, eds.), pp. 223–234, Springer International Publishing.

[7] N. Yamato, H. Sumioka, H. Ishiguro, M. Shiomi, and Y. Kohda, "Technology Acceptance Models from Different Viewpoints of Caregiver, Receiver, and Care Facility Administrator: Lessons from Long-Term Implementation Using Baby-Like Interactive Robot for Nursing Home Residents with Dementia," vol. 41, no. 4, pp. 296–321.

[8] J. Cohen-Mansfield, M. S. Marx, M. Dakheel-Ali, N. G. Regier, and K. Thein, "Can Persons With Dementia Be Engaged With Stimuli?," vol. 18, no. 4, pp. 351–362.

[9] L. Mathur, M. Spitale, H. Xi, J. Li, and M. J. Matarić, "Modeling user empathy elicited by a robot storyteller," *arXiv preprint arXiv:2107.14345*, 2021.

[10] F. Shi, N. Yamato, H. Ishiguro, M. Shiomi, and H. Sumioka, "Baby schema in human-robot physical interaction: Influence of baby likeness in a communication robot on caregiving behavior," *Computers in Human Behavior: Artificial Humans*, vol. 4, p. 100150, 2025.

[11] D. L. Murman, "The impact of age on cognition," *Seminars in Hearing*, vol. 36, no. 3, pp. 111–121, 2015.

[12] K. Nashiro, M. Sakaki, and M. Mather, "Age differences in brain activity during emotion processing: reflections of age-related decline or increased emotion regulation?," *Gerontology*, vol. 58, no. 2, pp. 156–163, 2012.

[13] "Heider's Theory of Balance:," vol. 21, no. 2, pp. 177–210.

[14] M. A. Harris and U. Orth, "The link between self-esteem and social relationships: A meta-analysis of longitudinal studies," *Journal of Personality and Social Psychology*, vol. 119, no. 6, pp. 1459–1477, 2020.

[15] A. Üstündağ, S. Haydaroğlu, D. Sayan, and M. Güngör, "The relationship between social anxiety levels and effective communication skills of adolescents participating in sports," *Scientific Reports*, vol. 15, no. 1, p. 15724, 2025.

[16] V. E. Caballo, G. B. Carrillo, and T. H. Ollendick, "Effectiveness of a social skills play-based training program intervention for childhood social anxiety," *Behavioral Psychology / Psicología Conductual*, vol. 23, no. 3, pp. 403–427, 2015.

[17] C. Okamura, "Confirmative examination of social self-esteem inventory with tsbi translated in japanese," *Bulletin of Tsukuba International Junior College*, vol. 33, pp. 141–152, 2005.

[18] A. Aron, E. Melinat, E. N. Aron, R. D. Vallone, and R. J. Bator, "The experimental generation of interpersonal closeness: A procedure and some preliminary findings," *Personality and Social Psychology Bulletin*, vol. 23, no. 4, pp. 363–377, 1997.

[19] R. J. Lewicki and B. B. Bunker, "Developing and maintaining trust in working relationships," in *Trust in Organizations: Frontiers of Theory and Research*

(R. M. Kramer and T. R. Tyler, eds.), pp. 114–139, Thousand Oaks, CA: Sage, 1996.

[20] A. Carpenter and K. Greene, "Social penetration theory," in *The International Encyclopedia of Interpersonal Communication* (C. R. Berger and M. E. Roloff, eds.), pp. 1–4, John Wiley & Sons, Ltd, 2015.

[21] R. J. Lewicki, "Developing and maintaining trust in work relationships," *Trust in organizations: Frontiers of theory and research/Sage*, 1996.

[22] C. Bartneck and J. Hu, "Exploring the Abuse of Robots," vol. 9, pp. 415–433.

[23] R. F. Baumeister and M. R. Leary, "The need to belong: Desire for interpersonal attachments as a fundamental human motivation," vol. 117, no. 3, pp. 497–529.

[24] C. Becker-Asano, T. Kanda, C. Ishi, and H. Ishiguro, "How about laughter? Perceived naturalness of two laughing humanoid robots," pp. 1–6.

[25] D. S. Berry and L. Z. McArthur, "Some components and consequences of a babyface," vol. 48, no. 2, pp. 312–323.

[26] J. I. D. Campbell and V. A. Thompson, "MorePower 6.0 for ANOVA with relational confidence intervals and Bayesian analysis," vol. 44, no. 4, pp. 1255–1265.

[27] N. Fairbrother, R. G. Barr, J. Pauwels, R. Brant, and J. Green, "Maternal thoughts of harm in response to infant crying: An experimental analysis," vol. 18, no. 3, pp. 447–455.

[28] S. Feng, N. Yamato, H. Ishiguro, M. Shiomi, and H. Sumioka, "Baby schema in human-robot physical interaction: Influence of baby likeness in a communication robot on caregiving behavior," p. 100150.

[29] S. Feng, N. Yamato, H. Ishiguro, M. Shiomi, and H. Sumioka, "Effect of Emotional Expression on the Impression of Older People Towards Baby-like Robots: Effect of Emotional Expression of Baby-like Robots on Older People," in *Proceedings of the 12th International Conference on Human-Agent Interaction*, pp. 414–416, Association for Computing Machinery.

[30] S. Feng, N. Yamato, H. Ishiguro, M. Shiomi, and H. Sumioka, "Effect of

Emotional Expression on the Impression of Older People Towards Baby-like Robots: Effect of Emotional Expression of Baby-like Robots on Older People," in *Proceedings of the 12th International Conference on Human-Agent Interaction*, pp. 414–416, Association for Computing Machinery.

[31] "The First Relationship."

[32] A. M. Frodi, M. E. Lamb, L. A. Leavitt, and W. L. Donovan, "Fathers' and mothers' responses to infant smiles and cries," vol. 1, pp. 187–198.

[33] Y. Furuta, M. Kanoh, T. Shimizu, M. Shimizu, and T. Nakamura, "Subjective evaluation of use of Babyloid for doll therapy,"

[34] Y. Furuta, M. Kanoh, T. Shimizu, M. Shimizu, and T. Nakamura, "Subjective evaluation of use of Babyloid for doll therapy," pp. 1–4.

[35] K. Hayashi, "Humanoid Robots as a Passive-Social Medium: A Field Experiment at a Train Station,"

[36] M. Heerink, B. Krose, V. Evers, and B. Wielinga, "Assessing Acceptance of Assistive Social Agent Technology by Older Adults: The Almere Model," vol. 2, pp. 361–375.

[37] J. Holler, K. H. Kendrick, M. Casillas, and S. C. Levinson, eds., *Turn-Taking in Human Communicative Interaction*. Frontiers Media SA.

[38] L. Hung, C. Liu, E. Woldum, A. Au-Yeung, A. Berndt, C. Wallsworth, N. Horne, M. Gregorio, J. Mann, and H. Chaudhury, "The benefits of and barriers to using a social robot PARO in care settings: A scoping review," vol. 19, no. 1, p. 232.

[39] K. Inoue, K. Wada, and T. Shibata, "Exploring the applicability of the robotic seal PARO to support caring for older persons with dementia within the home context," vol. 15, p. 26323524211030285.

[40] K. Inoue, K. Wada, and T. Shibata, "Exploring the applicability of the robotic seal PARO to support caring for older persons with dementia within the home context," vol. 15, p. 26323524211030285.

[41] C. A. Insko, "Balance Theory, The Jordan Paradigm, and The Wiest Tetrahedron," in *Advances in Experimental Social Psychology* (L. Berkowitz, ed.), vol. 18, pp. 89–140, Academic Press.

[42] F. Jimenez, M. Kanoh, and M. Goto, "How Do We Feel When Babyloid Starts Crying Suddenly?," in *Universal Access in Human-Computer Interaction. User and Context Diversity* (C. Stephanidis and M. Antona, eds.), pp. 355–362, Springer.

[43] M. Kanoh, "A Robot as"Receiver of Care" in Symbiosis with People,"

[44] J. Kim, S. Kim, S. Kim, E. Lee, Y. Heo, C.-Y. Hwang, Y.-Y. Choi, H.-J. Kong, H. Ryu, and H. Lee, "Companion robots for older adults: Rodgers' evolutionary concept analysis approach," vol. 14, no. 5, pp. 729–739.

[45] G. Livingston, A. Sommerlad, V. Orgeta, S. G. Costafreda, J. Huntley, D. Ames, C. Ballard, S. Banerjee, A. Burns, J. Cohen-Mansfield, C. Cooper, N. Fox, L. N. Gitlin, R. Howard, H. C. Kales, E. B. Larson, K. Ritchie, K. Rockwood, E. L. Sampson, Q. Samus, L. S. Schneider, G. Selbæk, L. Teri, and N. Mukadam, "Dementia prevention, intervention, and care," vol. 390, no. 10113, pp. 2673–2734.

[46] G. Livingston, J. Huntley, A. Sommerlad, D. Ames, C. Ballard, S. Banerjee, C. Brayne, A. Burns, J. Cohen-Mansfield, C. Cooper, S. G. Costafreda, A. Dias, N. Fox, L. N. Gitlin, R. Howard, H. C. Kales, M. Kivimäki, E. B. Larson, A. Ogunniyi, V. Orgeta, K. Ritchie, K. Rockwood, E. L. Sampson, Q. Samus, L. S. Schneider, G. Selbæk, L. Teri, and N. Mukadam, "Dementia prevention, intervention, and care: 2020 report of the Lancet Commission," vol. 396, no. 10248, pp. 413–446.

[47] M. B. Mathur and D. B. Reichling, "Navigating a social world with robot partners: A quantitative cartography of the Uncanny Valley," vol. 146, pp. 22–32.

[48] K. Nashiro, M. Sakaki, and M. Mather, "Age Differences in Brain Activity during Emotion Processing: Reflections of Age-Related Decline or Increased Emotion Regulation?," vol. 58, no. 2, pp. 156–163.

[49] H. Nittono and N. Ihara, "Psychophysiological Responses to Kawaii Pictures With or Without Baby Schema," vol. 7, no. 2, p. 2158244017709321.

[50] T. Obo and K. Takizawa, "Analysis of Timing and Effect of Visual Cue on Turn-Taking in Human-Robot Interaction," vol. 34, no. 1, pp. 55–63.

[51] M. Owczarek, E. Nolan, M. Shevlin, S. Butter, T. Karatzias, O. McBride, J. Murphy, F. Vallieres, R. Bentall, A. Martinez, and P. Hyland, "How is

loneliness related to anxiety and depression: A population-based network analysis in the early lockdown period," vol. 57, no. 5, pp. 585–596.

[52] T. Ozeki, T. Mouri, H. Sugiura, Y. Yano, and K. Miyosawa, "Use of communication robots to converse with people suffering from schizophrenia," vol. 7, no. 1, p. 13.

[53] "Relationship Between Loneliness, Psychiatric Disorders and Physical Health ? A Review on the Psychological Aspects of Loneliness - PMC."

[54] A. Rezzani, "Investigating Robot Abuse," in *Proceedings of the 2024 International Conference on Advanced Visual Interfaces*, pp. 1–3, Association for Computing Machinery.

[55] "A Robot as "Receiver of Care" in Symbiosis with People," vol. 27, no. 6, pp. 193–201.

[56] F. Yamaoka, T. Kanda, H. Ishiguro, and N. Hagita, "Interacting with a human or a humanoid robot?," in *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, (San Diego, CA, USA), pp. 2685–2691, IEEE, 2007.

[57] M. Rutherford and M. Przednowek, "Fathers show modifications of infant-directed action similar to that of mothers," vol. 111, pp. 367–78.

[58] D. Sakamoto, K. Hayashi, T. Kanda, M. Shiomi, S. Koizumi, H. Ishiguro, T. Ogasawara, and N. Hagita, "Humanoid Robots as a Broadcasting Communication Medium in Open Public Spaces," vol. 1, no. 2, pp. 157–169.

[59] K. Sander, Y. Frome, and H. Scheich, "FMRI activations of amygdala, cingulate cortex, and auditory cortex by infant laughing and crying," vol. 28, no. 10, pp. 1007–1022.

[60] M. Shiomi and H. Sumioka, "Differential effects of robot's touch on perceived emotions and the feeling of Kawaii in adults and seniors," vol. 15, no. 1, p. 7590.

[61] M. Shiomi, "A systematic survey of multiple social robots as a passive- and interactive-social medium," vol. 0, no. 0, pp. 1–15.

[62] M. Shiomi, R. Hayashi, and H. Nittono, "Is two cuter than one? Number and relationship effects on the feeling of kawaii toward social robots,"

[63] T. Shiwa, T. Kanda, M. Imai, H. Ishiguro, and N. Hagita, "How quickly should communication robots respond?," in *Proceedings of the 3rd ACM/IEEE International Conference on Human Robot Interaction*, pp. 153–160, Association for Computing Machinery.

[64] G. Skantze, "Turn-taking in Conversational Systems and Human-Robot Interaction: A Review," vol. 67, p. 101178.

[65] Y. Song, A. Luximon, and Y. Luximon, "The effect of facial features on facial anthropomorphic trustworthiness in social robots," vol. 94, p. 103420.

[66] Y. Song and Y. Luximon, "Trust in AI Agent: A Systematic Review of Facial Anthropomorphic Trustworthiness for Social Robot Design," vol. 20, no. 18, p. 5087.

[67] K. K. Steinnes, J. K. Blomster, B. Seibt, J. H. Zickfeld, and A. P. Fiske, "Too Cute for Words: Cuteness Evokes the Heartwarming Emotion of Kama Muta," vol. 10.

[68] "Subjective evaluation of use of Babyloid for doll therapy — IEEE Conference Publication — IEEE Xplore."

[69] H. Sumioka, N. Yamato, M. Shiomi, and H. Ishiguro, "A Minimal Design of a Human Infant Presence: A Case Study Toward Interactive Doll Therapy for Older Adults With Dementia," vol. 8.

[70] H. Sumioka, K. Koda, S. Nishio, T. Minato, and H. Ishiguro, "Revisiting ancient design of human form for communication avatar: Design considerations from chronological development of Dogū," in *2013 IEEE RO-MAN*, pp. 726–731.

[71] "Is two cuter than one? number and relationship effects on the feeling of kawaii toward social robots — PLOS ONE."

[72] "Is two cuter than one? number and relationship effects on the feeling of kawaii toward social robots - PubMed."

[73] N. Valtorta and B. Hanratty, "Loneliness, isolation and the health of older adults: Do we need a new research agenda?," vol. 105, no. 12, pp. 518–522.

[74] A. Weiss and C. Bartneck, "Meta analysis of the usage of the Godspeed Questionnaire Series," pp. 381–388.

[75] N. Yamato, H. Sumioka, H. Ishiguro, M. Shiomi, and Y. Kohda, "Technology Acceptance Models from Different Viewpoints of Caregiver, Receiver, and Care Facility Administrator: Lessons from Long-Term Implementation Using Baby-Like Interactive Robot for Nursing Home Residents with Dementia," vol. 41, no. 4, pp. 296–321.

[76] A. Yasuhara and T. Takehara, "Robots with tears can convey enhanced sadness and elicit support intentions," vol. 10, p. 1121624.

[77] D. M. Zeifman and I. St James-Roberts, "Parenting the Crying Infant," vol. 15, pp. 149–154.

LIST OF PUBLICATIONS

Journal publications

[1] Shi Feng, Nobuo Yamato, Hiroshi Ishiguro, Masahiro Shiomi, Hidenobu Sumioka, "*Baby schema in human-robot physical interaction: Influence of baby likeness in a communication robot on caregiving behavior*". Computers in Human Behavior: Artificial Humans, Volume 4, 2025, 100150, ISSN 2949-8821, https://doi.org/10.1016/j.chbah.2025.100150.

[2] Shi Feng, Nobuo Yamato, Hiroshi Ishiguro, Masahiro Shiomi, Hidenobu Sumioka, "*Towards the Emotion Design of Baby-like Robots for Older Adults: Do Crying Baby Robots Attract People Like Real Infants?*". Advanced Robotics, 10.1080/01691864.2025.2544156.

[3] Shi Feng, Nobuo Yamato, Hiroshi Ishiguro, Masahiro Shiomi, Hidenobu Sumioka, "*Age Differences in Responses to Infantile Features of Baby-Like Robots: Baby Schema and Caregiving Behavior in Older and Younger Adults* ". submitted to Computers in Human Behavior: Artificial Humans.

Conference publications

[1] Shi Feng, Nobuo Yamato, Hiroshi Ishiguro, Masahiro Shiomi, Hidenobu Sumioka, "*How does baby likeness affect people? -Investigation on the effect of baby like appearance and voice of robots-*". Information Processing Society of Japan, Interaction 2023, 1B-42.

[2] Shi Feng, Nobuo Yamato, Hiroshi Ishiguro, Masahiro Shiomi, Hidenobu Sumioka, "*Effect of Emotional Expression on the Impression of Older People Towards Baby-like Robots: Effect of Emotional Expression of Baby-like Robots on Older People*".

HAI '24: Proceedings of the 12th International Conference on Human-Agent Interaction, Pages 414 - 416, https://doi.org/10.1145/3687272.3690901.